A Newton-Raphson Method for Numerically Constructing Invariant Curves

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Wolfgang Marty

aus

Zürich, ZH

begutachtet von

Prof. Dr. Stefan Sauter

Prof. Dr. Bodo Werner

Zürich 2009

Die vorliegende Arbeit wurde von der Mathematischen-naturwissentlichen Fakultät der Universität Zürich auf Antrag von Prof. Dr. Stefan Sauter und Prof. Dr. Michel Chipot als Dissertation angenommen.

Abstract

This thesis is concerned with the numerical construction of simply closed invariant curves of maps defined on the plane. We develop and discuss a Newton-Raphson method that is based on solving a linear functional equation. By using formal power series analytic solutions are derived and conditions for the existence of a unique 2π-periodic continuous solution are established. In order to approximate this particular solution a basis of functions is introduced and an infinite system of linear equations for the coefficients of the basis is considered. We solve a sequence of finite subsystems with increasing dimension. By using B-splines and Fourier series an algorithm for approximating the invariant curve is derived. The algorithm is tested with explicitly given maps, followed by the application to the Van-der-Pol equation and the logistic map. The implementation is checked extensively and the efficiency of the method is illustrated.

Contents

Conventions

This thesis consists of 9 chapters. The chapters are divided into sections. (1.2.3) denotes formula (3) of Section 2 in Chapter 1. If we refer to formula (3) in Section 1.2 we only write (3) otherwise we use the full reference (1.2.3). Within the chapters, definitions, assumptions, theorems and examples are numerated continually, e.g. Theorem 2.1 refers to Theorem 1 in Chapter 2. In Section 2.2 and Chapter 5, algorithms are discussed. They are numerated continually.

The end of a theorem, corollary, lemma, algorithm or example is marked with $\Diamond$. Square brackets [ ] contain references. The details of the references are given at the end of the thesis.

$\mathbf{N}$ is the set of the positive integers, $\mathbf{N_0} = \mathbf{N} \cup \{ 0 \}$ is the set of the nonnegative integers, $\mathbf{Z}$ is the set of the integers, $\mathbf{Q}$ is the set of the rational numbers and $\mathbf{R^1}$ is the set of the real numbers. For $n \in \mathbf{N}$, $\mathbf{R^n}$ is the set of n tupels $(x_1, x_2, \ldots, x_n)$ with $x_i \in \mathbf{R^1}$, $1 \leq i \leq n$. $\mathbf{C}$ is the set of the complex numbers.

1. Introduction

This work considers a differentiable map

$$F: \mathbf{R}^2 \to \mathbf{R}^2$$

and a simply closed curve I in $\mathbf{R}^2$ being mapped onto itself by F. The curve I is referred to as *an invariant curve under the map F.* We are concerned with *the condition of invariance,*

$$F(I) = I,$$

where F(I) denotes the set of all images of the points of I.

Invariant curves are important for analysing the map F. For $x_0 \in I$ we have $F^n(x_0) \in I$, n = 1, 2, 3,….. We therefore examine the dynamics of F on I as n increases, and the discussion of F in the neighbourhood of I gives insight into the stability properties of F. Additionally, the study of differential equations frequently requires the investigation of invariant curves as the long-term behaviour of their solutions is explored.

Invariant curves have been studied extensively (see e.g. Aronson et al. [2, 1982], Kuczma et al. [16, 1990]). They are useful in various applications arising in, for example, physics, biology and economics. The work we present here is concerned with the numerical approximation of invariant curves, for which only a few algorithms have been discussed so far. *Newton-Raphson methods* have been considered by Broer et al. [3, 1997], Kevrekidis et al. [11, 1985], Nicolaisen [20, 1998] and Osinga [21, 1996]. They are based on linearising a given map and have superior convergence behaviour compared with fixed-point iterations. Our algorithm proceeds along the same lines as the one proposed by Nicolaisen: by studying the corresponding linear problem in detail we have considerably extended some of his results. The condition of invariance is expressed as a nonlinear functional equation for the invariant curve. *Functional equations* are relationships between values of a function at different arguments where no derivatives are involved.

The algorithm presented in this thesis is based on two steps: In *step 1* we linearize the functional equation arising from the condition of invariance. This leads to a linear functional equation for the Newton-Raphson correction. This equation has many solutions, but in order to construct the Newton-Raphson algorithm we need to identify a unique continuous 2π-periodic solution. We determine the precise degree of differentiability of this particular solution. This has direct implications for our algorithm. Nicolaisen considers only spline-functions. We find that in cases of low differentiability cubic B splines are appropriate: in other cases Fourier polynomials are more suitable. In *step 2* we approximate numerically the particular solution investigated in step 1. Our algorithm and the algorithm of Nicolaisen necessitate solving systems of linear equations. We investigate the structure of the considered matrices, and as a consequence use whichever is the most appropriate solver for the numerical experiments. We compare direct with iterative solvers. Additionally, the Newton-Raphson method requires the derivative of the current approximation of I. In contrast to Nicolaisen, we do not use any approximations by finite differences.

In order to analyse our Newton-Raphson algorithm we consider the condition of invariance as a map from the set of the continuous 2π-periodic real-valued functions into itself. By considering the supremum norm a Banach space is introduced. The Fréchet derivative of this map does not exist, and in addition we have no control over the second derivatives. Consequently, the convergence theorems on Newton-Raphson methods in Kantorovitch and Akilov [10, 1981] are not applicable, and further theoretical analysis of the Newton-Raphson method as developed in this thesis is left to future research. However, the results we obtain on the corresponding linear problem of the Newton-Raphson method are a useful starting point for studying the convergence of the algorithm presented here.

The numerical experiments illustrate that our algorithm computes smooth invariant curves successfully. We have numerically tested the condition of invariance, and by choosing different initial approximations for our algorithm the accuracy of the computed points on I has been examined.

We proceed with an overview of the literature that is concerned with the numerical approximation of smooth invariant curves: In Kirchgraber and Stiefel [12, 1978] invariant manifolds are analysed. The invariant curves we consider are one-dimensional invariant manifolds. The existence and uniqueness proof of invariant manifolds in Kirchgraber and Stiefel suggests a fixed-point iteration for computing invariant curves. However, as our work shows, in many cases the resulting algorithm produces a slow convergence. Nevertheless, we have used it as a comparison with the Newton-Raphson algorithm presented here.

Chang [4, 1983] discusses the rotation number of I under F. The rotation number describes the dynamics of the circle map induced by the diffeomorphism from the invariant curve to the unit circle. The algorithm proposed by Chang assumes that the rotation number of I is irrational. In this case truncated Fourier series can achieve high levels of accuracy. We also have illustrated by several numerical examples that a Fourier series approach can be very suitable for approximating invariant curves. However, we show that Fourier series may converge slowly if the rotation number of I is rational.

Van Veldhuizen [27, 1987], [28, 1988] approximates the invariant curve by polygons. The approach is based on an iteration scheme for the given map F. The convergence of the algorithm is discussed. The number of discretisation knots is fixed during the approximation process. The accuracy depends on the number of knots. In our method the required accuracy of the approximation of I is *given* a priori, and the algorithm automatically introduces the necessary number of knots.

The algorithm of Osinga [21, 1996] is based on discretising the condition of invariance and uses linear interpolation. Osinga's approach is not limited to invariant curves, and higher dimensional invariant manifolds are also considered.

As mentioned before, Newton-Raphson methods for computing smooth invariant curves have been proposed by Kevrekidis et al. [11, 1985] and Nicolaisen [20, 1998]. The algorithm of Kevrekidis is based on discretising the condition of invariance followed by solving the nonlinear equations by Newton-Raphson methods. On the other hand linearising the condition of invariance *first,* followed by numerically solving the corresponding linear problem is the approach used by Nicolaisen: in the work presented here we also adopt this method. Our algorithm is based on a specific functional equation that has been investigated by Kuczma [15, 1968]. Consequently, Kuczma's results are used as a starting point for this thesis.

1.1. The description of an invariant curve

We are concerned with a differentiable function F that is defined on $\mathbf{R}^2$ and maps $\mathbf{R}^2$ into $\mathbf{R}^2$, i.e.

$$F: x \in \mathbf{R}^2 \rightarrow \tilde{x} = F(x) \in \mathbf{R}^2 \qquad (1.1.1)$$

where $\tilde{x} \in \mathbf{R}^2$ denotes the corresponding image of $x \in \mathbf{R}^2$.

In the following the concept of *invariance* is discussed. Assuming that F has a fixed point $s \in \mathbf{R}^2$ satisfying

$$F(s) = s,$$

s *is invariant under F.*

Let I be a curve in $\mathbf{R}^2$ and F(I) denotes the image of I under F. We consider

$$F(I) \subseteq I, \qquad (1.1.2)$$

i.e. I is mapped by F into itself. Curves I satisfying (2) in the neighbourhood of a fixed point $s \in I$ of F are discussed in Kuczma [16]. Conditions for their existence and uniqueness are derived.

As illustrated in Sections 1.2 and 1.4 many maps F give rise to a simply closed curve I satisfying (2). The following definition introduces the curves considered in this thesis.

Definition 1.1: A simply closed, differentiable curve I is said to be *invariant under F* if the restriction of the map F to I satisfies the following conditions:

1. F is a map onto I, that is each point on I is the image of another point of I.

2. F is invertible on I.

It is seen from Definition 1.1 that we require that I is mapped by F onto I, i.e. not only (2) but

$$F(I) = I$$

is assumed. In this thesis special focus is given to the behaviour of invariant curves satisfying Definition 1.1 in the neighbourhood of a fixed point s of F.

Let $\| \quad \|_2$ denote the Euclidian norm in $\mathbf{R}^2$.

Definition 1.2: The *distance of* $x \in \mathbf{R}^2$ *from* I is defined by

$$\mathrm{dist}(x, I) = \min_{y \in I} \|x - y\|_2 .$$

We proceed with the description of the neighbourhood of I.

Definition 1.3: An *neighbourhood* $U(\varepsilon)$, $\varepsilon > 0$ *of* I is given by the set of points $x \in \mathbf{R}^2$ which satisfy

$$\mathrm{dist}(x, I) < \varepsilon.$$

Since I is simply closed the invariant curve can be parametrised by a function S(t), $t \in [0, T]$, $T > 0$ with $S(0) = S(T)$ and as I is assumed to be differentiable the orthogonal vector N(t) to the tangent vector in each point of I exists.

Assumption 1.1: There exists an neighbourhood $U(\varepsilon)$, $\varepsilon > 0$ of I such that $x \in U(\varepsilon)$ is uniquely represented by

$$x = S(t) + \zeta \cdot N(t),\ \zeta \in \mathbf{R}^1,\ t \in [0, T].$$

We note that each point x in $U(\varepsilon)$ is described by a specific value of t and, assuming $\|N(t)\|_2 = 1$, the orthogonal deviation ζ of x from I. Furthermore we have

$$\mathrm{dist}(x, I) = |\,\zeta\,|.$$

(t, ζ) are referred to as tubular coordinates [28, Veldhuizen].

Definition 1.4: I is called *attracting or repelling, respectively under F* if there exists an neighbourhood $U(\varepsilon)$, $\varepsilon > 0$ and a constant $0 < \kappa < 1$ such that

$$\mathrm{dist}(F(x), I) < \kappa\ \mathrm{dist}(x, I) \text{ or}$$

$$\text{dist}(x, I) < \kappa \, \text{dist}(F(x), I), \text{ respectively}$$

for all $x \in U(\varepsilon)$.

In this thesis we are specifically concerned with invariant curves I that satisfy properties 1 and 2 of the following assumption:

Assumption 1.2: There exists an neighbourhood $U(\varepsilon)$, $\varepsilon > 0$ of I such that

1. I is the only invariant curve under F in $U(\varepsilon)$.

2. I is either attracting or repelling under F.

In this thesis we are concerned with the numerical computation of attracting or repelling invariant curves I. More precisely, we want to compute arbitrary points on I with a specified a priori precision.

Considering the iteration

$$x_{n+1} = F(x_n),\ x_{n+1} \in U(\varepsilon),\ \varepsilon > 0,\ n = 0, 1, 2, \ldots.$$

is not sufficient for *computing arbitrary points on I* because

1. If I is attracting and contains an attracting fixed point satisfying $F(s) = s$ the iteration of F converges locally towards s. Arbitrary points on I are not approximated by the iteration (see Section 9.2).

2. If I is repelling the iteration does not converge towards points on I.

The examples below as well as the algorithms of this thesis are conveniently formulated in polar coordinates. By introducing an origin O in $\mathbf{R}^2$, r is the distance of $x \in \mathbf{R}^2$ from O and φ denotes the corresponding polar angle.

Assumption 1.3: We consider curves which delimit a starlike region relative to the origin. Hence there exists an origin for polar coordinates and a continuous function R such that the curve is represented by the equation

$$r = R(\varphi),\ \varphi \in \mathbf{R}^1$$

with

$$R(\varphi + 2\pi) = R(\varphi),\ \varphi \in \mathbf{R}^1. \tag{1.1.3}$$

With $r \in \mathbf{R}^1$, $\varphi \in \mathbf{R}^1$ as originals and $\tilde{r} \in \mathbf{R}^1$, $\tilde{\varphi} \in \mathbf{R}^1$ as corresponding images the function F defined by (1) is considered in polar coordinates and we express (1) by two continuous real-valued functions G and H of two variables r and φ:

$$\begin{pmatrix} r \\ \varphi \end{pmatrix} \in \mathbf{R}^2 \to \begin{pmatrix} \tilde{r} = G(r,\varphi) \\ \tilde{\varphi} = H(r,\varphi) \end{pmatrix} \in \mathbf{R}^2 \tag{1.1.4}$$

The use of polar coordinates implies the periodicity conditions

$$\begin{gathered} G(r, \varphi) = G(r, \varphi + 2\pi) \\ H(r, \varphi) + 2\pi = H(r, \varphi + 2\pi). \end{gathered} \tag{1.1.5}$$

(5) still holds if functions R satisfying Assumption 1.3 are substituted in (5):

$$\begin{gathered} G(R(\varphi), \varphi) = G(R(\varphi + 2\pi), \varphi + 2\pi) \\ H(R(\varphi), \varphi) + 2\pi = H(R(\varphi + 2\pi), \varphi + 2\pi). \end{gathered} \tag{1.1.6}$$

Let I be an invariant curve that satisfies Assumption 1.3. I is represented by the function $r = S(\varphi)$, $\varphi \in \mathbf{R}^1$. As it is assumed in Definition 1.1 that F maps I onto I it follows

$$\tilde{r} = S(\tilde{\varphi}),\ \tilde{\varphi} \in \mathbf{R}^1$$

for the image $\tilde{r}$ of r, $\tilde{\varphi}$ of φ, respectively. Substituting (4) and $r = S(\varphi)$ yield:

$$G(S(\varphi), \varphi) = S(H(S(\varphi), \varphi)). \tag{1.1.7}$$

In this relationship different arguments φ of S are connected with each other, that is we deal with a functional equation for the unknown function $r = S(\varphi)$.

Example 1.1: Following Veldhuizen [28] we consider

$$G(r, \varphi) = 1 + \kappa \cdot (r - 1)$$
$$H(r, \varphi) = \varphi + \tau \tag{1.1.8}$$

in (4). It is assumed that $\kappa \in \mathbf{R}^1$ and $\tau \in \mathbf{R}^1$ are given. The condition of invariance yields

$$S(\varphi + \tau) - \kappa \cdot S(\varphi) = 1 - \kappa$$

and it is seen that $S(\varphi) = 1$, $\varphi \in \mathbf{R}^1$ is a solution. Hence the unit circle is invariant under (8).

◊

Contrary to Example 1.1, many maps (4) give rise to invariant curves that cannot be determined explicitly. The method developed in this work starts with an initial approximation S_0 of the invariant curve followed by the computation of increasingly accurate approximations S_n, n = 1, 2, 3,…. of S. Our aim is to construct approximations $S(\varphi_1)$, $S(\varphi_2)$,…., $S(\varphi_N)$ where $\varphi_1, \varphi_2, \ldots, \varphi_N$ are apriori given polar angles.

As illustrated in Example 1.1, the evaluation of the map H in $\varphi_1, \varphi_2, \ldots, \varphi_N$ introduces in general for $\tau \neq 0$ polar angles different from $\varphi_1, \varphi_2, \ldots, \varphi_N$ and it is seen from Algorithm 1 (Section 2.2) that a numerical method which allows the interpolation of S_n, n = 0, 1, 2,…. in any given polar angle φ is requested. In this work Fourier series and B-splines are used. In addition we compare the two methods for representing S. Veldhuizen [28] represents the approximations of S by polygons. However, if the approximations S_n, n = 0, 1, 2,…. and S can be parametrised by polar coordinates, Fourier series are a natural choice because the required 2π-periodicity is inherent.

In the algorithm developed in Veldhuizen [28] the number of discretisation knots is fixed during the approximation process. The approximation is shown to improve as the number of knots increases. The accuracy depends of the number of knots. In our method the required accuracy of the approximation of S is *given a priori* and the algorithm automatically introduces the necessary number of knots.

1.2. Applications of an algorithm for computing invariant curve

1. The Van-der-Pol [26] equation is studied in the theory of ordinary differential equations. Invariant curves describe the region of attraction of an asymptotically stable equilibrium solution [12, Kirchgraber]. Each solution with initial condition within the invariant curve is bounded and converges towards the equilibrium solution. Invariant curves separate the bounded from the unbounded solutions.

2. In [19, Maynard Smith] mathematical modelling in biology is studied. The book is concerned with the dynamics of population growth.

Let the sequence $v_n \in \mathbf{R}^1$, n = 0, 1, 2,…. denote the population density in generation n and let α be a parameter reflecting the growth rate. Starting from an initial population density v_n, we consider the relationship

$$v_{n+1} = \alpha\, v_n,\ \alpha > 0.$$

If $\alpha < 1$ there follows

$$\lim_{n\to\infty} v_n = 0$$

and the population becomes extinct. If $\alpha > 1$ the sequence v_n is unbounded and the population grows without limit. It is more realistic to model the growth of a population whose ability to reproduce in any generation is governed by the population in the previous generation:

$$\alpha_n = \gamma\,(1 - v_{n-1}),\ \gamma \in \mathbf{R}^1,\ n = 1, 2,\ldots.$$

and

$$v_{n+1} = \alpha_n \cdot v_n.$$

For example, the reproduction of a herbivorous species will depend on the vegetation, which may in turn depend on how much of the vegetation was eaten by herbivores in the previous year. Assuming that initial values v_0, v_1 are given, the model of *the delayed regulation* is then defined by

$$v_{n+1} = \gamma\, v_n\, (1 - v_{n-1}). \tag{1.2.1}$$

The nonlinear term $\gamma\, v_{n-1}\, v_n$ regulating the population models a time delay of one generation.

Equation (1) can be transformed into a system of 2 difference equations by introducing the new variable $u_n = v_{n-1}$:

$$u_{n+1} = v_n$$

$$v_{n+1} = \gamma\, v_n\, (1 - u_n).$$

Thus we consider the map

$$F : x = \begin{pmatrix} u \\ v \end{pmatrix} \to F(x) = \begin{pmatrix} v \\ \gamma\, v (1 - u) \end{pmatrix}. \tag{1.2.2}$$

F has a fixed point at

$$u = v = \frac{\gamma - 1}{\gamma}$$

which is stable for $1 < \gamma \le 2$, i.e. the population density converges towards $\frac{\gamma}{\gamma - 1}$. For $\gamma > 2$ the fixed point becomes instable and there exists an attracting invariant curve that surrounds the fixed point. When γ reaches approximately 2.27 the invariant curve breaks down, leaving chaotic behaviour [2, Aronson].

1.3. Summary of this thesis

In Chapter 2 we introduce Newton-Raphson-type iterations for computing zeros of maps that are defined on a Banach space. P denotes the set of the continuous 2π-periodic real-valued functions of the variable $\varphi \in \mathbf{R}^1$ and the distance between $R \in P$ is defined by

$$\|R\|_{\sup} = \sup_{\varphi \in \mathbf{R}^1} |R(\varphi)|.$$

Then P is a Banach space. Assuming that G and H are given by (1.1.4) we are concerned with the map b defined by

$$b(R)(\varphi) = G(R(\varphi), \varphi) - R(H(R(\varphi), \varphi)).$$

From $R \in P$ and multiple use of (1.1.3) and (1.1.6) yields

$$b(R)(\varphi + 2\pi) = b(R)(\varphi).$$

As G and H are assumed to be continuous there follows $b(R) \in P$. Thus b is a map $P \rightarrow P$ and it is seen that solving (1.1.7) is equivalent to finding a zero of the map b.

In Section 2.2 we derive the algorithm discussed in this thesis. It is assumed that a zero S of b exists. Starting point is an initial approximation S_0 of S and we construct a sequence S_n, n = 0, 1, 2,…. of increasingly more accurate approximations of S by linearization of b in S_n. By using the composition sign $\circ$ the solution d_n of

$$L_{S_n}(d_n) = b_n \tag{1.3.1a}$$

with

$$L_{S_n}(d_n) = d_n \circ h_n - a_n \cdot d_n \tag{1.3.1b}$$

approximates the deviation of S_n to S. The coefficient functions h_n, a_n and b_n of (1) can be calculated from G, H and their partial derivatives with respect to r. By substituting $r = S_n(\varphi)$ it is seen that h_n, a_n and b_n are functions of φ. On the right side of (1b) different arguments of d_n are connected with each other, i.e. we deal with a functional equation for d_n. We discuss extensively the solutions of (1) (see Chapters 3 and 4). Assuming the periodicity conditions (1.1.5) and $|a_n(\varphi)| \neq 1$, $\forall \varphi \in \mathbf{R}^1$ it is shown in Theorem 2.1 that there exists a unique solution $d_n \in P$.

We consider

$$S_{n+1} = S_n + d_n, \; n = 0, 1, 2, \ldots. \tag{1.3.2}$$

where d_n is the solution of (1). As discussed in Chapter 2 the iteration (2) is a Newton-Raphson method. In this thesis we present realizations of (2) that numerically converge quadratically towards a zero S of the map b. An important property of the algorithm developed in this work is that it can be *applied for computing attracting and repelling invariant curves.* The convergence of Newton-Raphson methods for maps that are defined on a Banach space is investigated in [10, Kantorovich]. However, the results in [8] are not directly applicable to the problem considered in this thesis (see Section 2.1).

By suppressing S_n and the index n in (1) it is seen that in each iteration of the Newton-Raphson method

$$(f \circ h)(\varphi) - a(\varphi) \cdot f(\varphi) = b(\varphi), \; \varphi \in \mathbf{R}^1 \tag{1.3.3}$$

has to be solved for the unknown function f. The left side of (3) is linear in f, i.e. we are concerned with a *linear functional equation for f* and in Chapter 3 *the corresponding homogeneous linear functional equation*

$$(g \circ h)(\varphi) - a(\varphi) \cdot g(\varphi) = 0, \; \varphi \in \mathbf{R}^1 \tag{1.3.4}$$

is also considered.

We proceed with an overview of this thesis. In Sections 3.1, 3.2 and in [15, Kuczma] the solutions of (3) and (4) are examined in the case of h having one single fixed point. In 3.1 f is discussed by using formal power series. We derive recurrence formulae which express the coefficients of the formal power series for f, g in (3), (4), respectively, by the coefficients of the formal power series of h, a and b. For specific choices of h, a, b the theorems are illustrated (Example 3.1, 3.2, see also [30, Waldvogel]). In 3.2 it is shown that under the assumption that h, a, b are real-analytic the series introduced in Theorem 2.1 are converging in a complex neighbourhood of a fixed point of h. Thus we conclude that they represent a real-analytic solution f of (3). In Section 4.1 we expand the discussion of (3) from one fixed point to two fixed points of h.

In Sections 4.2 and 4.3 it is assume that h is a monotonically increasing function $\mathbf{R}^1 \to \mathbf{R}^1$ with $h(\varphi + 2\pi) = 2\pi + h(\varphi)$ [7, De Melo] and that a and b are 2π-periodic. We consider infinitely many times differentiable coefficient functions h, a, b and discuss in detail the regularity of the special solution $f \in P$ derived in Theorem 2.1. With denoting the k^{th} iterate of h by h^k the regularity depends critically on the uniquely determined rotation number of h given by

$$\rho = \frac{1}{2\pi} \lim_{k \to \infty} \frac{h^k(\varphi_0)}{k} \in \mathbf{R}^1, \ \varphi_0 \in \mathbf{R}^1. \tag{1.3.5}$$

If ρ is rational, the function f is only finitely many times differentiable. If ρ is irrational, we derive conditions under which f is infinitely many times differentiable.

Chapter 5 deals with the numerical computation of $f \in P$. However, the evaluation of the series presented in Theorem 2.1 is insufficient for solving (3) because they often converge too slowly. The overall quadratic convergence of the Newton-Raphson method is lost as the resulting algorithm for solving (3) converges only linearly. We propose that $f \in P$ is expanded in a system of basis functions. By inserting into (3) and equating coefficients, we find that the coefficients have to satisfy an infinite system of linear equations. In each step we need to approximate a system of linear equations with an infinite number of unknowns. By looking at finite subsystems we study the approximation of its solution in detail (see Section 5.3). By increasing their dimensions, the numerical experiments in Chapters 6-9 show that the sequence of solutions converges very fast for the realisations of (3) considered in this thesis. This method is used for approximating the solution $f \in P$ of (3).

Chapter 6 contains a study of test examples. The map F in (1.1.1) is explicitly given by the functions G and H introduced in (1.1.4). In Chapter 7 the forced Van-der-Pol oscillator is examined. We illustrate our method by choosing different parameters.

In Chapter 8 it is shown that the algorithm developed in Section 5.2 can also be used for computing the rotation number (5) of

$$h(\varphi) = \varphi + \alpha \sin \varphi, \ \varphi \in \mathbf{R}^1 \tag{1.3.6}$$

with $\alpha \in \mathbf{R}^1$ and $|\alpha| < 1$. Although we restrict our numerical experiments to the study of (6) our method allows to generalise to more than two nontrivial Fourier coefficients. A comparison with the methods in [29, Veldhuizen] and [17, MacKay] remains open for future research. However, it is shown that our method is considerably more efficient than the direct approximation of the limit based on definition (5).

Chapter 9 deals with (1.2.2). We treat the cases $\gamma = 2.10$, $\gamma = 2.11$ with Fourier polynomials. Our algorithm is shown to be also successful if an iterate of F has fixed points on the invariant curve.

1.4. Introductory example

In this section we consider realisations of the map (1.1.4). It is illustrated that the development of a numerical method is required for studying invariant curves of the map F. The invariant curves of the first two examples, however, can be determined explicitly. The map considered in Example 1.1 is decoupled, i.e. $G(r, \varphi) = G(r)$, $H(r, \varphi) = H(\varphi)$.

Example 1.1 (continuation): We again consider map (1.1.8) and illustrate the Newton-Raphson method by choosing a circle $r_0 = S_0(\varphi)$, $\varphi \in \mathbf{R}^1$, $r_0 > 0$ as an initial approximation S_0 for S. Applying Algorithm 1 (see Section 2.2) shows that

$$h_0(\varphi) = \varphi + \tau, \; a_0(\varphi) = \kappa, \; b_0(\varphi) = 1 - r_0 + \kappa\,(r_0 - 1)$$

in (1.3.1). From Theorem 2.1 there follows that the linear functional equation (1.3.1) with $n = 0$ has a unique solution $d_0 \in P$. We find

$$d_0(\varphi) = 1 - r_0$$

and (1.3.2) yields

$$S(\varphi) = d_0(\varphi) + S_0(\varphi) = 1,$$

thus we find convergence in one step. Choosing a circle as initial condition it is seen that the method developed in this thesis can easily deal with this map.

◊

In the following example we illustrate (1.1.7) by a map that can be handled by equating coefficients. The invariant curve is explicitly given by a Fourier series.

Example 1.2: Let

$$\begin{aligned} G(r, \varphi) &= \kappa\, r + T(\varphi) \\ H(r, \varphi) &= \varphi + \tau \end{aligned} \qquad (1.4.1)$$

in (1.1.4) where $\tau \in \mathbf{R}^1$, $\kappa \in \mathbf{R}^1$ with $|\kappa| \neq 1$ and T with $T(\varphi + 2\pi) = T(\varphi)$ are given. Furthermore let T be real-analytic with the Fourier series

$$T(\varphi) = \sum_{k=-\infty}^{\infty} t_k e^{ik\varphi}, \; t_k \in \mathbf{C}.$$

We show that for the map (1) there is a unique real-analytic 2π-periodic solution

$$S(\varphi) = \sum_{k=-\infty}^{\infty} S_k e^{ik\varphi}, \; S_k \in \mathbf{C} \tag{1.4.2}$$

to (1.1.7). Inserting of (2) into (1.1.7) and using (1) yield:

$$\sum_{k=-\infty}^{\infty} S_k e^{ik\varphi} \; (e^{ik\tau} - \kappa) = \sum_{k=-\infty}^{\infty} t_k e^{ik\varphi}$$

and by equating coefficients of $e^{ik\varphi}$ we find:

$$S_k = \frac{t_k}{e^{ik\tau} - \kappa}, \; k \in \mathbf{Z}. \tag{1.4.3}$$

Since $|\kappa| \neq 1$ by hypothesis there exists $c > 0$ such that $\left|e^{ik\tau} - \kappa\right| > c$, $k \in \mathbf{Z}$ and using (3), we have

$$|S_k| < \frac{|t_k|}{c}, \; k \in \mathbf{Z}. \tag{1.4.4}$$

As T is assumed to be real-analytic, it follows that $\overline{t_{-k}} = t_k$, $k \in \mathbf{Z}$ and from (3) we conclude $\overline{S_{-k}} = S_k$, $k \in \mathbf{Z}$. From (3) and (4), we see that (2) converges in a strip of the complex plane which contains the real axis. Therefore, the series (2) represents a real-analytic 2π-periodic function which fulfils (1.1.7). Hence the map (1) has at least one invariant curve.

To prove the uniqueness of S, assume by contradiction that there exists a second real-analytic 2π-periodic solution S of (1.1.7) for the map (1). We show that the function

$$d(\varphi) = \hat{S}(\varphi) - S(\varphi) = \sum_{k=-\infty}^{\infty} d_k e^{ik\varphi}, \; d_k \in \mathbf{C} \tag{1.4.5}$$

vanishes identically. (1.1.7) and (1) give:

$$\kappa\, S(\varphi) + T(\varphi) = S(\varphi + \tau)$$

$$\kappa\, \hat{S}(\varphi) + T(\varphi) = \hat{S}(\varphi + \tau),$$

thus

$$\kappa\, (S(\varphi) - \hat{S}(\varphi)) + \hat{S}(\varphi + \tau) - S(\varphi + \tau) = 0.$$

Using (5), we have:

$$\sum_{k=-\infty}^{\infty} d_k e^{ik\varphi} \quad (\kappa - e^{ik\tau}) = 0$$

and as the left side represents a real-analytic function that vanishes identically we find

$$d_k\, (\kappa - e^{ik\tau}) = 0,\ k \in \mathbf{Z}.$$

As

$$\kappa - e^{ik\tau} \neq 0,\ k \in \mathbf{Z}$$

by hypothesis, there follows

$$d_k = 0,\ k \in \mathbf{Z},$$

which implies $S(\varphi) = \hat{S}(\varphi)$, $\varphi \in \mathbf{R^1}$.

Theorem 2.1 of Section 2.3 implies also that map (1) has at most one invariant curve and consequently S is unique.

◊

In the study of ordinary differential equations invariant curves are of special interest. The map F is not explicitly given but has to be approximated by numerical integration of the given differential equation. F is called the *Poincaré map*. However, there is a close connection between maps F and ordinary differential equations [12, page 227, Kirchgraber]. Considering (1.1.4) in the form

$$\tilde{r} = \kappa\, r + T(r, \varphi),\ \kappa \in \mathbf{R^1}$$
$$\tilde{\varphi} = H(r, \varphi) \tag{1.4.6}$$

it is shown in [11] that (1.1.7) has a uniquely determined solution S if the constant $\kappa \in \mathbf{R^1}$ and the functions T and H satisfy various assumptions. The proof invokes Banach's fixed point theorem. In [12], however, the map (6) is studied for r, φ being vectors in $\mathbf{R^N}$. The discussion is thus not limited to 2 dimensions.

The map (6) is a generalisation of the map considered in Example 1.2. Contrary to (1), T and H are functions of the two variables r and φ. The condition of invariance (1.1.7) yields:

$$\kappa\, S(\varphi) + T(S(\varphi), \varphi) = S(H(S(\varphi), \varphi)).$$

Furthermore, assuming $\kappa \neq 0$, we have:

$$S = \frac{1}{\kappa}\,(S(H(S(\varphi), \varphi)) - T(S(\varphi), \varphi)).$$

Assuming that an initial approximation S_0 of S is given, we consider the iteration

$$S_{n+1} = \frac{1}{\kappa}\,(S_n(H(S_n(\varphi), \varphi)) - T(S_n(\varphi), \varphi)),\ n = 0, 1, 2,\ldots.. \tag{1.4.7}$$

If S_0 is sufficiently close to S and choosing $\kappa \geq 1.5$, we find that the iteration (7) converges to S for many maps H and T. In addition, we observe that the rate of convergence increases with increasing κ. The following example compares the iteration (7) and the Newton-Raphson method discussed in this work.

Example 1.3: Following (6), let

$$G(r, \varphi) = \kappa\, r + T(r, \varphi),\ \kappa \in \mathbf{R^1}$$
$$H(r, \varphi) = \varphi + p(r) \tag{1.4.8}$$

where $T(r, \varphi) = c\, r^2 + t(\varphi)$, $c \in \mathbf{R^1}$. It is assumed that t is a continuous 2π-periodic function and p is a polynomial in r. We choose the initial approximation $S_0(\varphi) = 0$ and represent the functions by real-valued Fourier polynomials. With

$$\kappa = 1.5,\ c = 0.1,\ t(\varphi) = 0.1 + \sin\varphi,\ p(r) = 1 + 0.4\,r + 0.1\,r^2 \qquad (1.4.9)$$

in (8), the iteration (7) converges in 5-digit (10-digit) precision with 33 (74) iterations and 2636 (10152) calls of map (8) are necessary. With the chosen data (9), the number of times Algorithm 1 in Section 2.2 needs to evaluate (8) is substantially less (see Section 6.1). The algorithm developed in this thesis is therefore particularly attractive for studying invariant curves of the Poincaré map because the computation of $\tilde{r}$ and $\tilde{\varphi}$ in (1.1.4) is not only the evaluation of a simple analytic expression but the numerical integration of an ordinary differential equation.

◊

As shown in Chapters 6-9, Algorithm 1 in this thesis is applicable to many maps in the form given by (1.1.4). The approximations S_n, $n = 1, 2, 3, \ldots$ of S have been examined extensively by numerical evaluation of the functional equation (1.1.7). In addition, we have chosen different initial approximations S_0 for testing the accuracy of the computed points on S.

2. A Newton-Raphson method for constructing invariant curves

2.1. Some notions from functional analysis

Following Section 1.3 we are concerned with the set P of the continuous 2π-periodic real-valued functions of the variable $\varphi \in \mathbf{R}^1$. Defining the norm $R \in P$ by

$$\|R\|_{sup} = \sup_{\varphi \in \mathbf{R}^1} |R(\varphi)| \qquad (2.1.1)$$

P becomes a Banach space. We use the notation $\| \ \|$ for $\| \ \|_{sup}$ throughout this chapter.

We are concerned with a map B defined on an open subset Ω of P:

$$B: \Omega \to P. \qquad (2.1.2)$$

Remark: The map B realised as residual (2.2.1a) of the condition of invariance (1.1.7) is in general not differentiable on P and consequently does not follow the general framework as described in the following. However, the iterations (5)-(8) in this section yield the starting point for investigating the nondifferentiability of B discussed e.g. in Chapter 4.

Let L(P) denote the set of the bounded linear maps from P into P. In the following definition of the differentiability of a map B is introduced (see e.g. Trubowitz [25, 1987]).

Definition 2.1: A map B is *differentiable in $R \in \Omega$*, if there exists a $L_R \in L(P)$ such that

$$\|B(R+\Delta) - B(R) - L_R(\Delta)\| = o(\|\Delta\|), \ \|\Delta\| \to 0,$$

i.e. for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|B(R+\Delta) - B(R) - L_R(\Delta)\| \le \varepsilon \|\Delta\| \qquad (2.1.3)$$

for all Δ with $\| \Delta \| < \delta$, $\Delta \in \Omega$. The linear map L_R is uniquely determined, and is called *the first derivative* of B at R. The map B is *differentiable on Ω*, if it is differentiable at each point R in Ω.

Definition 2.2: If B: $\Omega \rightarrow P$ is differentiable on Ω and if its derivative L_R: $\Omega \rightarrow L(P)$ is also differentiable on Ω, then B is called *twice differentiable*. Its second derivative is a map

$$L_R^{(2)} : \Omega \rightarrow L^{(2)}(P),\ R \in \Omega$$

from Ω into the Banach space $L^{(2)}(P)$ of all bounded, bilinear maps from P x P into P.

Let $S \in \Omega$ be a zero of the map (2):

$$B(S) = 0.$$

Furthermore let S_0 be an approximation of S. We consider

$$B(S) - B(S_0) + B(S_0) = 0. \tag{2.1.4}$$

For formulating Newton-Raphson type iterations for approximating S we need the following:

Assumption 2.1: The map B is differentiable in $R \in \Omega$ and the first derivative L_R of B is invertible in R.

(Original Newton-Raphson process): Substituting $R = S_0$ and $\Delta = \Delta_0$ with $\Delta_0 = S - S_0$ in (3) and approximating $B(S) - B(S_0)$ by $L_{S_0}(\Delta)$ in (4) yield:

$$L_{S_0}(d_0) + B(S_0) = 0$$

where d_0 is unknown and it follows:

$$d_0 = -\left(L_{S_0}\right)^{-1}(B(S_0)).$$

The first step of the iteration

$$S_1 = S_0 + d_0$$

leads to

$$S_{n+1} = S_n + d_n,\ n = 0, 1, 2, \ldots \tag{2.1.5a}$$

where

$$d_n = -\left(L_{S_n}\right)^{-1}\left(B\left(S_n\right)\right). \tag{2.1.5b}$$

(Theoretical Newton-Raphson process): Theoretically assuming that the zero S of the map B is known we also consider the substitution R = S and $\Delta = \Delta_0$ with $\Delta_0 = S_0 - S$ in (3) and approximating $B(S) - B(S_0)$ by $L_S(\Delta)$ in (4) yields:

$$L_S(\hat{d}_0) + B(S_0) = 0,$$

where $\hat{d}_0$ is unknown and it follows:

$$\hat{d}_0 = -\left(L_S\right)^{-1}\left(B\left(S_0\right)\right).$$

The first step of the iteration

$$\hat{S}_1 = S_0 + \hat{d}_0$$

leads with $\hat{S}_0 = S_0$ to

$$\hat{S}_{n+1} = \hat{S}_n + \hat{d}_n,\ n = 0, 1, 2, \ldots \tag{2.1.6a}$$

where

$$\hat{d}_n = -\left(L_S\right)^{-1}\left(B\left(\hat{S}_n\right)\right). \tag{2.1.6b}$$

(Distance from the invariant curve): The iterations (5) and (6) are based on linear approximation by neglecting higher terms. By taking higher order terms into consideration we find:

$$S_n - S = \left(L_{S_n}\right)^{-1}\left(B\left(S_n\right) + O\left(\left(S - S_n\right)^2\right)\right),\ n = 0, 1, 2, \ldots \tag{2.1.7}$$

(Simplified Newton-Raphson method): With $S_0^* = S_0$, we also consider

$$S_{n+1}^* = S_n^* + d_n^* \text{ , } n = 0, 1, 2, \ldots \tag{2.1.8a}$$

where

$$d_n^* = -\left(L_{S_0}\right)^{-1}\left(B\left(S_n^*\right)\right). \tag{2.1.8b}$$

Contrary to the original Newton-Raphson process (5) the linear map L has to be inverted only once in (8) but the convergence speed of (5) is faster than (8) (see paragraph D in Section 6.2 for a numerical example).

The convergence of the sequences (5) and (8) is investigated in Kantorovich [10, 1981] assuming that a given differentiable operator

1. maps a given Banach space into itself.

2. has an invertible linear approximation.

However, as mentioned before, the operator B applied to invariant curves is in general not differentiable and as a consequence the application of the theorems in Kantorovich is not obvious. Thus we need a more specific analysis for studying Newton-type iterations applied to maps in the plane. The investigation of the sequence (5) realised with maps (1.1.1) is part of this thesis.

2.2. The basic algorithm

In this section we apply the Newton-Raphson method (2.1.5) to maps (1.1.4). The basic algorithm of this thesis is formulated. It computes simple closed curves which are invariant under maps given by (1.1.4). Following (1.1.7) we consider the residual

$$b(R)(\varphi) = G(R(\varphi), \varphi) - R(H(R(\varphi), \varphi)). \qquad (2.2.1a)$$

Based on Assumption 1.3, we are only concerned with simply closed curves which delimit a starlike region relative to the origin. Such curves can be parametrised by polar coordinates. In order to reflect the 2π-periodicity properties of polar coordinates we consider the set P^m, m = 0, 1, 2,…. of the real-valued, m-times continuously differentiable, 2π-periodic function of a variable $\varphi \in \mathbf{R}^1$. P^0 is denoted with P and we use the sup-norm defined by (2.1.1) throughout the discussion in this section.

Assumption 2.2: Let the functions G and H given by (1.1.4) be continuous maps $\mathbf{R}^2 \to \mathbf{R}^1$ that are twice partially differentiable with respect to r. Furthermore the periodicity conditions (1.1.5) are assumed.

Let G_r, F_r, G_{rr}, F_{rr}, R', R'', Δ', $\tilde{\Delta}'$ denote the first and the second partial derivatives with respect to r, φ respectively. Taylor expansion and multiple application of (1.1.6) yield:

Lemma 2.1: Let G and H satisfy Assumption 2.2. Then the following holds:

1. If $R \in P^1$, the first derivative L_R of –b is given by

$$L_R(\Delta)(\varphi) = \Delta(H(R(\varphi), \varphi)) -$$

$$\{(G_r(R(\varphi), \varphi) - R'(H(R(\varphi), \varphi)) \cdot H_r(R(\varphi), \varphi)\} \cdot \Delta(\varphi), \Delta \in P \qquad (2.2.1b)$$

with $L_R(\Delta) \in P$.

2. If $R \in P^2$, the second derivative $L_R^{(2)}$ of –b is given by

$L_R^{(2)}(\Delta\tilde{\Delta})(\varphi) = \{\Delta'(H(R(\varphi), \varphi)) \cdot \tilde{\Delta}(\varphi) -$

$\Delta(\varphi) \cdot \tilde{\Delta}'(H(R(\varphi), \varphi))\} \cdot H_r(R(\varphi), \varphi) -$

$$\{G_{rr}(R(\varphi), \varphi) - (R''(H(R(\varphi), \varphi)) \cdot (H_r(R(\varphi), \varphi))^2 - (R'(H(R(\varphi),\varphi))) \cdot \quad (2.2.1c)$$

$(H_{rr}(R(\varphi), \varphi))\} \cdot \Delta(\varphi) \cdot \tilde{\Delta}(\varphi), \Delta(\varphi), \tilde{\Delta}(\varphi) \in P^1$

with $L_R^{(2)}(\Delta, \tilde{\Delta}) \in P$.

◊

With (1.1.7) it follows that an invariant curve represented by $r = S(\varphi)$ has to satisfy

$$G(S(\varphi), \varphi) - S(H(S(\varphi), \varphi)) = 0.$$

Thus zeros S of the map b defined by (1a) have to be determined. We apply the Newton-Raphson method formulated in Section 2.1 to maps (1.1.4). Assuming that an initial approximation S_0 of S is given, we substitute $R = S_n$, $\Delta = d_n$, respectively in (1a), (1b), respectively and (2.1.5) shows that in each Newton-Raphson step

$$L_{S_n}(d_n) = b(S_n),\ n = 0, 1, 2,\ldots.$$

has to be solved for d_n. Thus, by (1a) and (1b), a Newton-Raphson method for maps $\mathbf{R}^2 \to \mathbf{R}^2$ is defined.

The following algorithm *assumes* that the Newton-Raphson approximations are *differentiable* and is an initial attempt for computing invariant curves:

Algorithm 1

Given: 1. maps G and H satisfying Assumption 2.2;
2. initial approximation $S_0 \in P$;
3. ε: requested precision of the computed invariant curve.

BEGIN
$n = 0$;

REPEAT
1. Solve

$$(d_n \circ h_n)(\varphi) - a_n(\varphi) \cdot d_n(\varphi) = b_n(\varphi) \tag{2.2.2a}$$

for the unknown function d_n where

$$h_n(\varphi) = h(S_n)(\varphi) = H(S_n(\varphi), \varphi), \tag{2.2.2b}$$

$$a_n(\varphi) = a(S_n)(\varphi) = G_r(S_n(\varphi), \varphi) - S_n'(H(S_n(\varphi), \varphi)) \cdot H_r(S_n(\varphi), \varphi), \tag{2.2.2c}$$

$$b_n(\varphi) = b(S_n)(\varphi) = -S_n(H(S_n(\varphi), \varphi)) + G(S_n(\varphi), \varphi). \tag{2.2.2d}$$

2. Compute

$$S_{n+1}(\varphi) = S_n(\varphi) + d_n(\varphi). \tag{2.2.3}$$

3. Let $n = n + 1$;

UNTIL $\|d_n\| < \varepsilon$;

END.

◊

2.3. The 2π-periodic solution

If we assume $S_n \in P^1$ for n = 0, 1 , 2,…. it follows with (1.1.6) that a_n in (2.2.2c) and b_n in (2.2.2d) are 2π-periodic. In addition, we have

$$h_n(\varphi) = \varphi + p_n(\varphi) \tag{2.3.1}$$

with $p_n(\varphi + 2\pi) = p_n(\varphi)$, $\varphi \in \mathbf{R^1}$ in (2.2.2b). h_n is called *a circle map* [7, De Melo]. As we are investigating a single Newton-Raphson step in the following, we simplify the notation by suppressing the index n in the formulae (2.2.2). In each Newton-Raphson step we have to solve the *linear functional equation*

$$f \circ h - a \cdot f = b \tag{2.3.2}$$

for the unknown function f. h, a and b are called *the coefficient functions* of (2). We introduce the linear map

$$L(f) = f \circ h - a \cdot f \tag{2.3.3}$$

which, with (2) yields the equation

$$L(f) = b.$$

The *iterates* h^k, $k \in \mathbf{N}$ of h are given by

$$h^{k+l} = h \circ h^k \text{ for } k = 0, 1, 2,…. \tag{2.3.4}$$

where

$$h^0(\varphi) = \varphi.$$

Furthermore we define the null product as

$$\prod_{k=n}^{n-1} c_k = 1,\ c_k \in \mathbf{R^1},\ n \in \mathbf{N_0}. \tag{2.3.5}$$

Theorem 2.1 below is important for this work since it ensures the existence and uniqueness of a solution $f \in P$ of (2). In addition, it yields a series representation for the solution f of (2).

Theorem 2.1: Let $h(\varphi) = \varphi + p(\varphi)$, $p \in P$, $a \in P$ with $|a(\varphi)| \neq 1$, $\forall \varphi \in \mathbf{R^1}$ and $b \in P$ in the functional equation (2). We distinguish two cases:

1. If $|a(\varphi)| > 1$, $\forall \varphi \in \mathbf{R^1}$, we define

$$f = -\sum_{n=0}^{\infty} \frac{b \circ h^n}{\prod_{k=0}^{n} a \circ h^k}. \tag{2.3.6}$$

2. It is assumed that h is invertible and the inverse h^{-1} of h is denoted by $\tilde{h}$. If $|a(\varphi)| < 1$, $\forall \varphi \in \mathbf{R^1}$, we define

$$f = \sum_{n=0}^{\infty} b \circ \tilde{h}^{n+1} \prod_{k=1}^{n} a \circ \tilde{h}^k. \tag{2.3.7}$$

In both cases (2) has a unique solution $f \in P$. f is represented by the convergent series (6), (7), respectively.

Proof: First we consider case 1. Let $A_n = \prod_{k=1}^{n} \left(a \circ h^k\right)^{-1}$, $n = 0, 1, 2, \ldots$.

with $A_0(\varphi) = 1$, $\forall \varphi \in \mathbf{R^1}$. The series (6) satisfies (2) formally:

$$f \circ h = -\sum_{n=1}^{\infty} b \circ h^n \cdot A_n = -\sum_{n=0}^{\infty} b \circ h^n \cdot A_n + b = a \cdot f + b. \tag{2.3.8}$$

(1) implies:

$$h(\varphi + 2\pi) = h(\varphi) + 2\pi. \tag{2.3.9}$$

By using (4) induction with respect to n yields:

$$h^n(\varphi + 2\pi) = h^n(\varphi) + 2\pi,\ n = 0, 1, 2, \ldots. \tag{2.3.10}$$

From (6) and $a \in P$, $b \in P$ it follows:

$$f(\varphi + 2\pi) = -\sum_{n=0}^{\infty} \frac{b(h^n(\varphi + 2\pi))}{\prod_{k=0}^{n} a(h^k(\varphi + 2\pi))} = f(\varphi). \tag{2.3.11}$$

With b = R in (2.1.1) we find

$$|b(\varphi)| \leq \|b\|, \ \forall \varphi \ \in \mathbf{R^1}. \tag{2.3.12}$$

As we have assumed $|a(\varphi)| > 1$, $\forall \varphi \in \mathbf{R^1}$ with $a \in P$ and since $[0, 2\pi]$ is compact, there exists $C_a \in \mathbf{R^1}$ with

$$|a(\varphi)| > C_a > 1.$$

By the triangle inequality, we have:

$$|f| \leq \sum_{n=0}^{\infty} \frac{|b \circ h^n|}{\prod_{k=0}^{n} |a \circ h^k|} \leq \|b\| \left(\frac{1}{C_a} + \frac{1}{(C_a)^2} + ... \right) = \frac{\|b\|}{C_a - 1} < \infty . \tag{2.3.13}$$

The convergence of (6) is therefore uniform and it follows that f is continuous. Together with (11) the existence of a solution $f \in P$ is shown.

To prove uniqueness, let there exist a second solution $g \in P$ of (2). The difference of the two solutions satisfies:

$$|f - g| = \frac{1}{|a|} |f \circ h - g \circ h|.$$

Consider the $(n-1)^{st}$ iterate:

$$|f - g| = \frac{1}{|a|} .... \frac{1}{|a \circ h^{n-1}|} |f \circ h^n - g \circ h^n|, \ n = 0, 1, 2, ... \tag{2.3.14}$$

From $f \in P$, $g \in P$ it follows that $|f \circ h^n - g \circ h^n|$, n = 0, 1, 2,…. is bounded by a real constant, independent of n. With

$$|a(\varphi)| > 1, \ \forall \varphi \in \mathbf{R^1}$$

we have

$$\prod_{k=0}^{n-1} \frac{1}{\left|a \circ h^k\right|} \to 0 \ \text{ for } n \to \infty .$$

By using (14), we conclude:

$$\left|f - g\right| \to 0 \ \text{ for } n \to \infty$$

and we find

$$f = g. \tag{2.3.15}$$

The assertion of case 1 is thus demonstrated.

Consider case 2. By hypothesis $\tilde{h} = h^{-1}$ exists and as in (8), we find with

$$\tilde{A}_n = \prod_{k=0}^{n-1} a \circ \tilde{h}^k, \ n = 0, 1, 2, \ldots .$$

that

$$f \circ h = \sum_{n=0}^{\infty} b \circ \tilde{h}^n \cdot \tilde{A}_n = \sum_{n=0}^{\infty} b \circ \tilde{h}^{n+1} \cdot \tilde{A}_n + b = a \cdot f + b, \tag{2.3.16}$$

i.e. (7) satisfies (2). By applying (9) to $\tilde{\varphi} = h^{-1}(\varphi)$, $\varphi \in \mathbf{R}^1$, we have

$$h^{-1}(\varphi + 2\pi) = h^{-1}(\varphi) + 2\pi, \ \varphi \in \mathbf{R}^1. \tag{2.3.17}$$

Using (1) yields

$$h^{-1}(\varphi) = \varphi + p_1(\varphi), \tag{2.3.18}$$

where $p_1 = -p \circ h^{-1} \in P$. Because of (17), h can be replaced by $\tilde{h} = h^{-1}$ in (10) and it is seen that (10) is valid for $n \in \mathbf{Z}$. As in (11) we find that

$$f(\varphi + 2\pi) = \sum_{n=0}^{\infty} b(\tilde{h}^{n+1}(\varphi + 2\pi)) \prod_{k=1}^{n} a(\tilde{h}^k(\varphi + 2\pi)) = f(\varphi),$$

i.e. f is 2π-periodic. By hypothesis $\left|a(\varphi)\right| < 1$, $\forall \varphi \in \mathbf{R}^1$ and it follows that

$$\exists\, C_a \in \mathbf{R^1} \text{ with } |a(\varphi)| < C_a < 1,\ \varphi \in \mathbf{R^1}.$$

With (12), (13) we have:

$$|f| \le \sum_{n=0}^{\infty} \left|b \circ h^{-n-1}\right| \prod_{k=1}^{n} \left|a \circ h^{-k}\right| \le \|b\|(1 + C_a + (C_a)^2 + ...) = \frac{\|b\|}{1 - C_a}. \quad (2.3.19)$$

The conclusions from (13) can also be applied to (19) and it follows that

$$f \in P.$$

To prove uniqueness of $f \in P$, let $g \in P$ be a second solution of (2). As we have assumed that h is an invertible function $\mathbf{R^1} \to \mathbf{R^1}$, the function h^{-n} exists as well. We replace in (14) h^n by h^{-n}, $n = 1, 2, 3, \ldots$ and solve for $|f - g|$:

$$|f - g| = \left|a \circ h^{-1}\right| ...... \left|a \circ h^{-n}\right| \left|f \circ h^{-n} - g \circ h^{-n}\right|,\ n = 0, 1, 2, \ldots \quad (2.3.20)$$

By considering $n \to \infty$ and using the assumption $|a(\varphi)| < 1$, $\forall \varphi \in \mathbf{R^1}$, we find (15). This concludes the proof of Theorem 2.1.

◊

Corollary 1.1: Let the assumptions of Theorem 2.1 hold. If

$$|a(\varphi)| > 0,\ \forall \varphi \in \mathbf{R^1} \quad (2.3.21)$$

the assertion in case 2 of Theorem 2.1 is a consequence of the assertion in case 1.

Proof: Using (2) and (21) we have

$$f - \frac{f \circ h}{a} = -\frac{b}{a}.$$

This is the same as

$$f(\varphi) - \frac{f(h(\varphi))}{a(\varphi)} = -\frac{b(\varphi)}{a(\varphi)}. \quad (2.3.22)$$

As it is assumed in case 2 of Theorem 2.1 that h^{-1} exists φ can be replaced by $\tilde{h}(\varphi) = h^{-1}(\varphi)$ and we consider the linear functional equation

$$f \circ \tilde{h} - \tilde{a} \cdot f = \tilde{b} \tag{2.3.23}$$

where

$$\tilde{a} = \frac{1}{a \circ \tilde{h}}, \; \tilde{b} = -\frac{b \circ \tilde{h}}{a \circ \tilde{h}}. \tag{2.3.24}$$

With $a \in P$, $b \in P$ and (18) we have

$$\tilde{a} \in P, \; \tilde{b} \in P.$$

Using the assumption $|a(\varphi)| < 1, \; \forall \varphi \in \mathbf{R}^1$ of case 1 in Theorem 2.1 there exists $C \in \mathbf{R}^1$ with $0 < C < 1$ and

$$\left| \frac{1}{a \circ \tilde{h}} \right| > \frac{1}{C} > 1.$$

With (17) and (24) it follows that the linear functional equation (22) satisfies the assumptions of Theorem 2.1 in case 1 and consequently (22) has a unique solution $f \in P$. Furthermore f is represented by

$$f = -\sum_{n=0}^{\infty} \frac{\tilde{b} \circ \tilde{h}^n}{\prod_{k=0}^{n} \tilde{a} \circ \tilde{h}^k}.$$

The transition from (22) to (2) by substituting φ by $h(\varphi)$ and using the definition of $\tilde{a}$ and $\tilde{b}$ in (23) imply that (2) has a unique solution $f \in P$ represented by (7).

◊

3. The linear subproblem of the Newton-Raphson method

3.1. Formal power series

Following (2.3.2) we consider again the *linear functional equation*

$$f(h(x)) - a(x) \cdot f(x) = b(x). \tag{3.1.1}$$

In this section the coefficient functions h, a, b are given by formal power series expressed as powers of the variable x. The theorems below describe solutions f of (1) that are represented by formal power series. Let $\circ$ denote the composition of two formal power series. Thus (1) is the same as

$$(f \circ h)(x) - a(x) \cdot f(x) = b(x). \tag{3.1.2}$$

In Sections 3.1 and 3.2 it is assumed that h has a fixed point h(s) = s, $s \in \mathbf{R}^l$. With

$$x = s + y$$

in (2) it follows

$$(f \circ h)(s + y) - a(s + y) \cdot f(s + y) = b(s + y).$$

By defining

$$\tilde{f}(y) = f(s + y) \tag{3.1.3a}$$

we have

$$(\tilde{f} \circ \tilde{h})(y) - \tilde{a}(y) \cdot \tilde{f}(y) = \tilde{b}(y) \tag{3.1.3b}$$

where

$$\tilde{h}(y) = h(s + y) - s, \ \tilde{a}(y) = a(s + y), \ \tilde{b}(y) = b(s + y). \tag{3.1.3c}$$

We see that the coefficients

$$\tilde{f}(y) = \sum_{n=0}^{\infty} f_n y^n$$

in the fixed point $\widetilde{h}(0) = 0$ in (3c) are the same as the coefficients

$$f(s+y) = \sum_{n=0}^{\infty} f_n (s+y)^n$$

in the fixed point h(s) = s, s $\in \mathbf{R}^1$ in (2). For the discussion of f_n, n = 0, 1, 2,…. in s it is thus sufficient to assume h(0) = 0.

In the theorems below if n < j we use the convention that

$$\sum_{i=j}^{n} c_j = 0,\ c_j \in \mathbf{R}^1.$$

Theorem 3.1: In (2) let

$$h(x) = x \sum_{n=0}^{\infty} h_n x^n,\ h_n \in \mathbf{R}^1,\ a(x) = \sum_{n=0}^{\infty} a_n x^n,\ a_n \in \mathbf{R}^1,$$

$$b(x) = \sum_{n=0}^{\infty} b_n x^n,\ b_n \in \mathbf{R}^1 \tag{3.1.4}$$

and

$$a_0 \neq (h_0)^n,\ n = 0, 1, 2,\ldots.. \tag{3.1.5}$$

Then, there exists a unique formal power series

$$f(x) = \sum_{n=0}^{\infty} f_n x^n \tag{3.1.6}$$

that satisfies (2), and the coefficients f_n, n = 0, 1, 2,…. are recursively given by

$$f_n = \frac{b_n + \sum_{i=0}^{n-1} f_i a_{n-i} - \sum_{i=1}^{n-1} f_i p_{i,n-i}}{(h_0)^n - a_0} \tag{3.1.7}$$

where the coefficients $p_{n,i}$ are defined by

$$\left(\frac{h(x)}{x}\right)^n = p_n(x) = \sum_{i=0}^{\infty} p_{n,i} x^i, \; n = 0, 1, 2, \ldots. \qquad (3.1.8)$$

Proof: Using (4) and (6) in (2) and applying (8) yields:

$$\sum_{n=0}^{\infty} f_n x^n p_n(x) - \sum_{n=0}^{\infty} a_n x^n \cdot \sum_{n=0}^{\infty} f_n x^n = \sum_{n=0}^{\infty} b_n x^n. \qquad (3.1.9)$$

Equating coefficients for x^n, n = 0, 1, 2,…. shows that $f_i \cdot p_{i,n-i}$ and $f_i \cdot a_{n-i}$, i = 0, 1, 2,….,n contribute to x^n. Summation over i yields:

$$\sum_{i=0}^{n} f_i p_{i,n-i} - \sum_{i=0}^{n} f_i a_{n-i} = b_n, \; n = 0, 1, 2, \ldots.. \qquad (3.1.10)$$

From (8) it follows that

$$p_{n,0} = (h_0)^n, \; n = 0, 1, 2, \ldots., \; p_{0,i} = \delta_{0,i}, \; i = 0, 1, 2, \ldots., \qquad (3.1.11)$$

where $\delta_{k,n}$ is the Kronecker delta. With (10) we have:

$$f_n\,((h_0)^n - a_0) = b_n + \sum_{i=0}^{n-1} f_i a_{n-i} - \sum_{i=1}^{n-1} f_i p_{i,n-i}, \; n = 0, 1, 2, \ldots. \qquad (3.1.12)$$

and, using assumption (5), we find that f_n has the unique representation (7).

◊

In the following let $h_0 > 0$, $h_0 \neq 1$, $a_0 > 0$. Furthermore the real number μ is defined by

$$\mu = \frac{\log a_0}{\log h_0}. \qquad (3.1.13)$$

This is the same as

$$h_0^{\mu} = a_0. \qquad (3.1.14)$$

Condition (5) is equivalent to $\mu \notin \mathbf{N_0}$. If $\mu \in \mathbf{N_0}$ a denominator in (7) vanishes and Theorem 3.1 is not applicable. The case $\mu \in \mathbf{N_0}$ is discussed in Theorem 3.3. First we assume b(x) = 0 and consider the *homogeneous linear functional equation*

$$(g \circ h)(x) - a(x) \cdot g(x) = 0 \qquad (3.1.15)$$

for the unknown function g. The solutions of (15) are only determined up to a multiplicative constant $g_0 \in \mathbf{R^1}$. The application of Theorem 3.1 with $b_n = 0$, n = 0, 1, 2,…. in (7) yields $f_n = 0$, n = 0, 1, 2,…., i.e. the formal power series (6) is trivial. The following theorem introduces a nontrivial power series for the solution g of (15).

Theorem 3.2: Let

$$h(x) = x \sum_{n=0}^{\infty} h_n x^n,\ h_n \in \mathbf{R^1},\ h_0 > 0,\ h_0 \neq 1,$$

$$a(x) = \sum_{n=0}^{\infty} a_n x^n,\ a_n \in \mathbf{R^1},\ a_0 > 0 \qquad (3.1.16)$$

in the homogeneous functional equation (15). Then, there exists a formal power series

$$g(x) = x^{\mu}\,(g_0 + g_1 x + g_2 x^2 + \ldots.) \qquad (3.1.17)$$

for the solution g of (15) where μ is defined by (13). The coefficients in (17) are given by

$$g_n = \frac{\sum_{i=0}^{n-1} g_i (a_{n-i} - p_{i+\mu,n-i})}{(h_0)^{n+\mu} - a_0},\ n = 1, 2, \ldots.,\ g_0 \in \mathbf{R^1} \qquad (3.1.18)$$

where the coefficients $p_{n,i}$ are defined as in (8).

Proof: We use (16) and (17) in (15) and divide by x^{μ}:

$$\left(\frac{h(x)}{x}\right)^{\mu} \sum_{n=0}^{\infty} g_n (h(x))^n - \sum_{n=0}^{\infty} a_n x^n \cdot \sum_{n=0}^{\infty} g_n x^n = 0.$$

With (8) we have

$$\sum_{n=0}^{\infty} g_n x^n p_{n+\mu}(x) - \sum_{n=0}^{\infty} a_n x^n \cdot \sum_{n=0}^{\infty} g_n x^n = 0.$$

With $f_n = g_n$ and $b_n = 0$, n = 0, 1, 2,…. we see that this equation has the same form as (9). Using (10) and equating coefficient for x^n, n = 0, 1, 2,…. yield

$$\sum_{i=0}^{n} g_i p_{i+\mu,n-i} - \sum_{i=0}^{n} g_i a_{n-i} = 0, \; n = 0, 1, 2,\ldots.$$

thus we find

$$g_n\,(p_{n+\mu,0} - a_0) = \sum_{i=0}^{n-1} g_i a_{n-i} - \sum_{i=0}^{n-1} g_i p_{i+\mu,n-i}, \; n = 0, 1, 2,\ldots. \tag{3.1.19}$$

From (11) it follows

$$p_{n+\mu} = h_0^{n+\mu}, \; n = 0, 1, 2,\ldots, \tag{3.1.20}$$

and with $h_0 \neq 1$ we find

$$h_0^{n+\mu} \neq h_0^{\mu}, \; n = 1, 2,\ldots.$$

Using (14) und (20) we conclude that (18) follows from (19).

◊

Example 3.1: Let

$$h(x) = 2 \cdot x - x^2, \; a(x) = 2 \tag{3.1.21}$$

in (15). By evaluating (13) we find

$$\mu = 1.$$

Theorem 3.2 yields for n = 1, 2, 3:

$$g_1 = \frac{g_0}{2}, \; g_2 = \frac{g_0}{3}, \; g_3 = \frac{g_0}{4}.$$

The conjecture

$$g(x) = g_0 \sum_{n=0}^{\infty} \frac{x^n}{n} = g_0 \log(1 - x),\ g_0 \in \mathbf{R}^1 \tag{3.1.22}$$

is confirmed by inserting in (15). Thus, the series g defined by (22) is a one-parametric solution of (15) whereby the coefficient functions are given by (21).

◊

We consider the inhomogeneous linear functional equation (2) and the corresponding homogeneous equation (15) ($b(x) = 0$). It is assumed that the formal power series for (2) and (5) exists. From the linearity of (2) it follows that the formal series of

$$f(x) + g(x)$$

satisfies (2). We distinguish two cases:

1. $\mu > 0$
There exists a one-parametric family of solution g of (2). Considering $\mu \notin \mathbf{N_0}$ and assuming $g_0 \neq 0$ Theorem 3.2 shows that the powers of x are not integer. If $g_0 = 0$ it follows from Theorem 3.1 that the powers of x are integer, i.e. there exists a unique Taylor series (6) that satisfies (2).

2. $\mu < 0$
If $g_0 \neq 0$ the formal power series of g is singular in the origin. However, Theorem 3.1 shows that there exists a unique Taylor series (5) that satisfies (2).

In both case there exists a unique Taylor series. We discuss this special solution of (2) in more detail in Section 3.2.

Theorem 3.1 excludes

$$\frac{\log a_0}{\log h_0} \in \mathbf{N_0}.$$

In Theorem 3.2, however, the case $\mu \in \mathbf{N_0}$ is included. In

$$f(x) = \gamma\, g(x) \log x + E(x),\ \gamma \in \mathbf{R^1} \qquad (3.1.23)$$

the solution f of (2) is related to the solution g of the corresponding homogeneous functional equation (15). Substituting (23) in (2) yields:

$$\gamma\,(g \circ h)(x) \log h(x) + (E \circ h)(x) - a(x)\,[\gamma\, g(x) \log x + E(x)] = b(x),\ \gamma \in \mathbf{R^1}.$$

As g satisfies (15), it follows that

$$(E \circ h)\,(x) - a(x)\,E(x) = b(x) - \gamma\, a(x)\, g(x) \log \frac{h(x)}{x},\ \gamma \in \mathbf{R^1}. \qquad (3.1.24)$$

We note that (24) is a linear functional equation of the form (2) for the unknown formal power series

$$E(x) = \sum_{n=0}^{\infty} e_n x^n\ . \qquad (3.1.25)$$

In the following theorem, we normalise by $g_0 = 1$ in (17) and determine γ such that a Taylor series (25) exists which satisfies (23).

Theorem 3.3: The following structure is assumed in the linear functional equation (2):

1. Let

$$\begin{aligned} h(x) &= x \sum_{n=0}^{\infty} h_n x^n,\ h_n \in \mathbf{R^1},\ h_0 > 0,\ h_0 \neq 1, \\ a(x) &= \sum_{n=0}^{\infty} a_n x^n,\ a_n \in \mathbf{R^1},\ a_0 > 0,\ b(x) = \sum_{n=0}^{\infty} b_n x^n,\ b_n \in \mathbf{R^1} \end{aligned} \qquad (3.1.26)$$

where the formal power series of $\left(\frac{h(x)}{x}\right)^n$, n = 0, 1, 2,…. is given by (8).

2. Let $\mu = \dfrac{\log a_0}{\log h_0} \in \mathbf{N_0}$.

Then the following holds:
There exists a formal power series (25) for E(x) in (23). We consider three cases:

1. If $0 \le n < \mu$ we have:

$$e_n = \frac{b_n + \sum_{i=0}^{n-1} e_i a_{n-i} - \sum_{i=1}^{n-1} e_i p_{i,n-i}}{(h_0)^n - a_0}. \tag{3.1.27}$$

2. With

$$\gamma = \frac{b_\mu + \sum_{i=0}^{\mu-1} e_i a_{\mu-i} - \sum_{i=1}^{\mu-1} e_i p_{i,\mu-i}}{a_0 \log h_0} \tag{3.1.28}$$

in (25), the coefficient $e_\mu \in \mathbf{R}^1$ is arbitrary.

3. If $n > \mu$ we have:

$$e_n = \frac{b_n - d_n + \sum_{i=0}^{n-1} e_i a_{n-i} - \sum_{i=1}^{n-1} e_i p_{i,n-i}}{(h_0)^n - a_0}. \tag{3.1.29}$$

With

$$d_\mu = \gamma\, a_0 \log h_0$$

the coefficients $d_{\mu+1}, d_{\mu+2}, d_{\mu+3}, \ldots$ are defined by

$$\sum_{n=\mu}^{\infty} d_n x^n = \gamma\, a(x)\, g(x) \log \frac{h(x)}{x}, \tag{3.1.30}$$

where γ is given by (28) and g is the solution of the corresponding homogeneous linear functional (15) with $g_0 = 1$ of (2).

Proof: As $\mu \in \mathbf{N_0}$ by hypothesis it exists a $n \in \mathbf{N_0}$ such that (14) holds and the assumptions $h_0 > 0$, $h_0 \neq 1$ in (26) imply

$$a_0 \neq (h_0)^n, \ n \neq \mu. \tag{3.1.31}$$

As (24) is a linear functional equation of the form (2) for the unknown series (25) we follow the proof of Theorem 3.1 with $f_n = e_n$, $n = 0, 1, 2, \ldots$ and $n \neq \mu$.

Three cases are distinguished:

1. Let $0 \le n < \mu$.
The formal power series of g begins with x^{μ} (Theorem 3.2), i.e. the term

$$\gamma\, a(x)\, g(x) \log \frac{h(x)}{x}$$

in (25) does not influence the coefficient of x^n. Using (31), it follows that equation (12) can be solved for e_n and we find (27).

2. Let $n = \mu$.
From (15) with $g_0 = 1$ and (26) it follows $d_\mu = \gamma\, a_0 \log h_0$ in (30). Considering (10) we have with (12):

$$e_\mu\, p_{\mu,0} - e_\mu\, a_0 = b_\mu - \gamma\, a_0 \log h_0 + \sum_{i=0}^{\mu-1} e_i a_{\mu-i} - \sum_{i=1}^{\mu-1} e_i p_{i,\mu-i}\,. \qquad (3.1.32)$$

From (11) and (14) we conclude that the left-hand side of (32) vanishes. In addition $e_\mu \in \mathbf{R}^1$ is an arbitrary real number. The assumptions $a_0 > 0$ and $h_0 \neq 1$ yield

$$a_0 \log h_0 \neq 0.$$

The right-hand side of (32) can be solved for γ and we find (28).

3. Let $n > \mu$.
Using the notation (30) we see that $b_n - d_n$ is the n^{th} coefficient of the formal power series of

$$b(x) - \gamma\, a(x)\, g(x) \log \frac{h(x)}{x}\,.$$

With (31) it follows that (12) can be solved for e_n and we find (29).

◊

From Theorem 3.3 it follows that the series (25) is only determined up to a real constant $e_\mu \in \mathbf{R}^1$. With the choice $e_\mu = 0$ in (25) und (23) we obtain a special solution f of the inhomgeneous linear functional equation (2). In the following corollary we write the general solution of (2) as a sum of this special solution and the general solution (17) of the homogeneous functional equation (15).

Corollary 3.1: Let the assumptions of Theorem 3.3 be satisfied. Furthermore let

$$E^*(x) = \sum_{\substack{n=0 \\ n \neq \mu}}^{\infty} e_n x^n$$

where the coefficients e_n, $n = 1,\ldots, \mu - 1, \mu + 1,\ldots.$ are given by (27), (29) respectively. Then the solution f of (2) has the representation

$$f(x) = \gamma\, g(x) \log x + E^*(x) + e_\mu\, g(x),\ e_\mu \in \mathbf{R}^1, \tag{3.1.33}$$

where γ is given by (28) and g denotes the solution of the corresponding homogeneous linear functional equation (15) of (2).

◊

The following example illustrates Corollary 3.1.

Example 3.2: Let

$$h(x) = \frac{2x}{1+x},\ a(x) = 2,\ b(x) = 1 - x \tag{3.1.34}$$

in (2). From (13) we have

$$\mu = 1.$$

We consider the corresponding homogeneous problem (15) of (2) where the coefficient functions h and a are given by (34). With $g_0 = 1$ we find by evaluating (18)

$$g_1 = 1,\ g_2 = 1,\ g_3 = 1.$$

The conjecture

$$g(x) = \frac{x}{1-x}$$

is verified by evaluating (15) using (34). By (28), we have

$$\gamma = -\frac{1}{\log 4}$$

and (23) yields

$$f(x) = -\frac{1}{\log 4}\frac{x}{1-x}\log x + E(x).$$

Based on (27), (29) the coefficients e_n, n = 0, 1, 2,.... of the formal power series (25) can be calculated. For n = 0, 1, 2 we have

$$e_0 = -1,$$

$$e_1 \in \mathbf{R^1},$$

$$e_2 = e_1 + \frac{1}{2} - \frac{1}{\log 4}$$

and (33) gives

$$f(x) = -\frac{1}{\log 4}\frac{x}{1-x}\log x - 1 + \left(\frac{1}{2} - \frac{1}{\log 4}\right)x^2 + \ldots\ldots + e_1\frac{x}{1-x},\ e_1 \in \mathbf{R^1}.$$

◊

3.2. Real-analytic solutions

In this section we investigate real-analytic solutions f of (2.3.2) in a neighbourhood of the origin O of the complex plane **C** and, similar to Section 3.1, h(O) = O is assumed.

Furthermore let U_δ denote the interior of a circle centred in O with radius $\delta > 0$. The coefficient functions h, a, b of (2.3.2) are assumed to be real-analytic functions of a complex variable x in O, i.e. there exists a $\delta > 0$ such that the representation

$$h(x) = x\sum_{n=0}^{\infty} h_n x^n,\ h_n \in \mathbf{R}^1,\ a(x) = \sum_{n=0}^{\infty} a_n x^n,\ a_n \in \mathbf{R}^1, \qquad (3.2.1a)$$

$$b(x) = \sum_{n=0}^{\infty} b_n x^n,\ b_n \in \mathbf{R}^1,\ x \in U_\delta \qquad (3.2.1b)$$

holds.

Theorem 3.4: The following is assumed for the coefficient functions of the linear functional equation (2.3.2):

1. h, a, b are real-analytic functions in O represented by the series (1);

2. $h_0 \neq 0, |h_0| \neq 1, a_0 \neq 0$;

3. $a_0 \neq (h_0)^n$, n = 0, 1, 2,…. (condition (3.1.5)).

Then, there exists a unique real-analytic solution f of (2.3.2) in a neighbourhood of O.

We start with proving the following lemma:

Lemma 3.1: Let the assumptions of Theorem 3.4 be satisfied. Then the assertion of Theorem 3.4 follows under the additional assumption

$$|h_0| < 1. \qquad (3.2.2)$$

Proof: We start by demonstrating the *existence* of a real-analytic solution f. As the assumptions of Theorem 3.4 imply the assumptions of Theorem 3.1

there exists a formal power series (3.1.6) with (3.1.7) that satisfies (2.3.2). Let

$$P(x) = \sum_{k=0}^{m-1} f_k x^k, \; m \in \mathbf{N}, \; x \in U_\delta, \; \delta > 0. \tag{3.2.3}$$

Then

$$f^*(x) = f(x) - P(x) = \sum_{k=m}^{\infty} f_k x^k \tag{3.2.4}$$

is a formal power series satisfying

$$f^*(h(x)) - a(x) \cdot f^*(x) = b^*(x) \tag{3.2.5}$$

where

$$b^*(x) = b(x) + a(x) \cdot P(x) - P(h(x)).$$

As P satisfies (2.3.2) up to the power x^{m-1}, the power series of b* starts with x^m, thus $\exists\ C_1 > 0$ and $\exists\ \delta_1 \in \mathbf{R^1}$ with

$$0 < \delta_1 < 1 \tag{3.2.6}$$

such that

$$\left| b^*(x) \right| < C_1 \left| x \right|^m, \; x \in U_{\delta_1}. \tag{3.2.7}$$

As a is real-analytic and $a_0 = a(O) \neq 0$ by hypothesis of Theorem 3.4, $\exists\ C_2 > 0$ and $\exists\ \delta_2 \in \mathbf{R^1}$, $0 < \delta_2 < \delta_1$ with

$$\left| a(x) \right| > C_2, \; x \in U_{\delta_2}. \tag{3.2.8}$$

Let $q \in \mathbf{R^1}$ with $0 < q < 1$ be given. Assumption (2) implies the existence of $m^* \in \mathbf{N}$ with

$$\left| h_0 \right|^m < C_2\, q, \; m \geq m^*.$$

Then $\exists\ \delta_3 \in \mathbf{R^1}$, $0 < \delta_3 < \delta_2$ with

$$\left| h(x) \right|^m < C_2\, q \left| x \right|^m, \; x \in U_{\delta_3}, \; m \geq m^*. \tag{3.2.9}$$

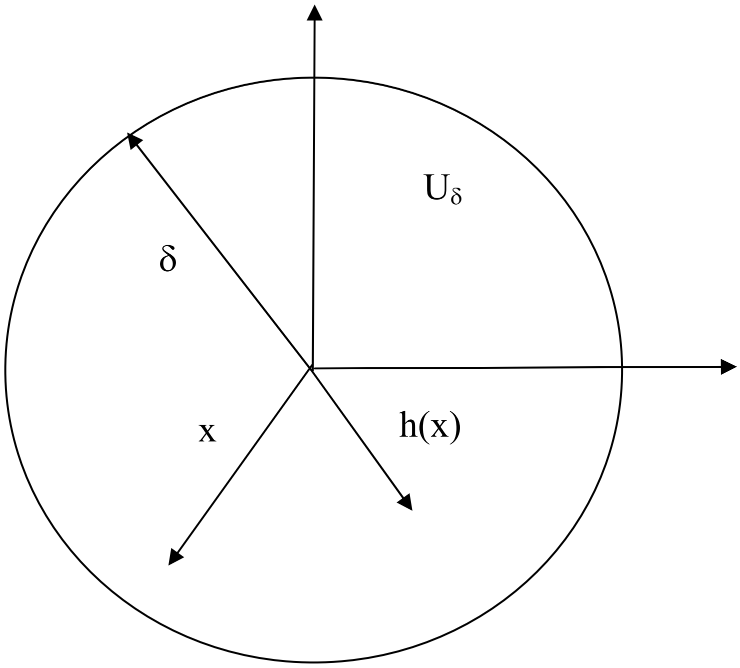

Figure 1

From (2) it follows that $\exists\delta_4 \in \mathbf{R}^1$, $0 < \delta_4 < \delta_3$ with

$$|h(x)| < |x|,\ x \in U_{\delta_4}$$

and

$$h(x) \in U_{\delta_4}$$

(see Figure 1). Iteration of (9) in U_{δ_4} yields:

$$\left|h^n(x)\right|^m < C_2^n\, q^n |x|^m,\ x \in U_{\delta_4},\ n = 0, 1, 2, \ldots \qquad (3.2.10)$$

With $f(x) = f^*(x)$, $x \in U_{\delta_4}$ in (2.3.6) we conclude that

$$f^*(x) = -\sum_{n=0}^{\infty} \frac{(b^* \circ h^n)(x)}{\prod_{k=0}^{n} (a \circ h^k)(x)} \qquad (3.2.11)$$

satisfies (5) formally. Using (7), (8) and (10) we have

$$\frac{\left|(b^* \circ h^n)(x)\right|}{\prod_{k=0}^{n}\left|(a \circ h^k)(x)\right|} \le C_1 \left|h^n(x)\right|^m C_2^{-n-1} < \tag{3.2.12}$$

$$C_1 C_2^{-1} q^n |x|^m, \; x \in U_{\delta_4}$$

for n = 0, 1, 2,…. and m ≥ m*. Let

$$C = C_1 C_2^{-1} \sup_{x \in U_{\delta_4}} |x|^m .$$

From (6) we have $0 < \delta_4 < 1$ and thus $C < C_1 C_2^{-1}$. Equation (11) and the summation over n in (12) give:

$$\left|f^*(x)\right| \le \frac{C}{1-q}, \; x \in U_{\delta_4} .$$

It is concluded that the series (11) converges independently of $x \in U_{\delta_4}$, i.e. the convergence is locally uniform. As it is assumed that the coefficient functions are real-analytic, it follows that f* and, using (4), f are real-analytic. The existence of a real-analytic solution f of (2.3.2) is thus demonstrated.

In order to show uniqueness we consider a second solution $\hat{f}$ of (2.3.2) in U_{δ_4} and show $\hat{f} = f$. The difference

$$g = f - \hat{f}$$

is real-analytic in U_{δ_4} and

$$g(x) = \sum_{n=0}^{\infty} g_n x^n, \; x \in U_{\delta_4}$$

satisfies the homogenous linear functional equation (3.1.15) where the coefficient functions h and a fulfil the assumptions of Theorem 3.4. The application of Theorem 3.1 with $f_n = g_n$ and $b_n = 0$, n = 0, 1, 2,…. yields $g_n = 0$, n = 0, 1, 2,…. As $g = f - \hat{f}$ is real-analytic in U_{δ_4} it follows g(x) = 0 and consequently $\hat{f}(x) = f(x)$, $x \in U_{\delta_4}$. The proof of Lemma 3.1 is thus concluded.

◊

Proof of Theorem 3.4: As it is assumed that the function a is real-analytic and $a_0 \neq 0 \;\exists\; C > 0$ with

$$0 < |a_0| < C$$

hence

$$\frac{1}{|a_0|} > \frac{1}{C} > 0$$

thus $\exists\; \delta_1 > 0$ with

$$\frac{1}{|a(x)|} > \frac{1}{C} > 0\,,\; x \in U_{\delta_1}\,. \tag{3.2.13}$$

Lemma 3.1 proves Theorem 3.4 under the additional assumption (2). As $|h_0| \neq 1$ is excluded by hypothesis it remains to consider $|h_0| > 1$. As $|h_0| \neq 0$, $\exists\; \delta_2 \in \mathbf{R}^1$, $0 < \delta_2 < \delta_1$ such that the real-analytic function $\tilde{h} = h^{-1}$ exists in U_{δ_2} and from h(O) = O we have $\tilde{h}(O) =$ O. Furthermore $\tilde{h}$ is represented by

$$\tilde{h}(x) = x \sum_{n=0}^{\infty} \tilde{h}_n x^n\,,\; \tilde{h}_n \in \mathbf{R}^1 \tag{3.2.14}$$

where $|\tilde{h}_0| < 1$ and consequently $\exists\, \delta_3 \in \mathbf{R}^1$, $0 < \delta_3 < \delta_2$ with

$$\tilde{h}(x) \in U_{\delta_2}\,,\; x \in U_{\delta_2}\,.$$

With (13) we have

$$\left|\frac{1}{(a \circ \tilde{h})(x)}\right| > \frac{1}{C} > 0,\; x \in U_{\delta_3}\,. \tag{3.2.15}$$

As a and b are real-analytic in the origin, $\exists\, \delta_4 \in \mathbf{R}^1$, $0 < \delta_4 < \delta_3$, such that

$$\tilde{a}(x) = \frac{1}{(a \circ \tilde{h})(x)},\; \tilde{b}(x) = -\frac{(b \circ \tilde{h})(x)}{(a \circ \tilde{h})(x)}$$

with

$$(f \circ \tilde{h})(x) - \tilde{a}(x) \cdot f(x) = \tilde{b}(x),\; x \in U_{\delta_3} \tag{3.2.16}$$

holds, i.e. $\tilde{a}$ and $\tilde{h}$ are real-analytic in O. From (15), we have that the function a does not vanish in O and with (14), it follows that (16) satisfies the assumptions of Lemma 3.1, i.e. there exists a unique real-analytic solution f of (16) in O. As $\tilde{h} = h^{-1}$ maps U_{δ_3} in a neighbourhood of O that contains U_{δ_4}, f remains real-analytic by replacing x by h(x) in (16). Thus, the transition from (16) to (2.3.2) shows that f is a unique real-analytic solution of (2.3.2) in O.

◊

In the following we consider the case $\mu \in \mathbf{N_0}$. From the proof of Theorem 3.4, Lemma 3.1, respectively it is seen that in order to show the existence of a real-analytic solution we only require the existence but not the uniqueness of the formal power series (3.1.6) and consequently of the polynomial (3). In the theorems below we apply Theorem 3.3 to (3.1.24) and since equating coefficients for the series (3.1.25) in Theorem 3.3 is not unique we need the following corollary:

Corollary 3.2: Let h, a, b in (2.3.2) be real-analytic in O with h(O) = O, $\left|h_0\right| \neq 1$, $h_0 \neq 0$, $a_0 \neq 0$. From the existence of a formal power series (3.1.6) it follows that there exists a real-analytic solution f of (2.3.2) in the neighbourhood of O.

◊

The case $\mu \in \mathbf{N_0}$ is first discussed for the homogeneous problem (3.1.15). We have:

Theorem 3.5: The following is assumed for the coefficient functions of the homogeneous linear functional equation (2.3.2):
1. h, a are real-analytic functions in O represented by the series in (1a);
2. h(O) = O, $h_0 > 0$, $h_0 \neq 1$, $a_0 > 0$;
3.

$$\mu = \frac{\log a_0}{\log h_0} \in \mathbf{N_0}. \tag{3.2.17}$$

Then, the homogeneous equation (3.1.15) has a nontrivial one-parametric real-analytic family of solutions g in the origin O. Furthermore g(O) = O and the order of the zero is μ.

First we prove the following lemma.

Lemma 3.2: Let the assumption of Theorem 3.5 be satisfied. Then the assertion of Theorem 3.5 follows under the additional assumption

$$0 < h_0 < 1. \tag{3.2.18}$$

Proof: First we assume

$$\mu = 0. \tag{3.2.19}$$

With (17) we have

$$a_0 = 1. \tag{3.2.20}$$

As the function a is real-analytic, $\exists\ \delta_1 \in \mathbf{R^1}$, $0 < \delta_1 < 1$ and (18) it follows that $\exists\ a_n \in \mathbf{R^1}$, $n = 1, 2, \ldots$ such that

$$a(x) = 1 + a_1 x + a_2 x^2 + \ldots,\ x \in U_{\delta_1}.$$

Then $\exists\ \delta_2 \in \mathbf{R^1}$, $0 < \delta_2 < \delta_1$ and $\exists\ C_1 > 0$ with

$$\log |a(x)| \leq C_1 |x|,\ x \in U_{\delta_2}, \tag{3.2.21}$$

Based on (18), $\exists\ \delta_3 \in \mathbf{R^1}$, $0 < \delta_3 < \delta_2$ and $\exists\ q \in \mathbf{R^1}$, $0 < q < 1$ with

$$|h(x)| \leq q |x|,\ x \in U_{\delta_3}$$

and

$$h(x) \in U_{\delta_3}$$

(see Figure 1). Iteration yields

$$\left|h^n(x)\right| \leq q^n |x|,\ x \in U_{\delta_3}. \tag{3.2.22}$$

The product

$$g(x) = g_0 \prod_{n=0}^{\infty} (a \circ h^n)(x))^{-1}, \; g_0 \in \mathbf{R^1} \tag{3.2.23}$$

satisfies (3.1.15) formally for $x \in U_{\delta_3}$. From h(O) = O and (19) we have

$$g(O) = g_0. \tag{3.2.24}$$

Furthermore using (18) and (20), $\exists\, \delta_4 \in \mathbf{R^1}$, $0 < \delta_4 < \delta_3$ and a function $\hat{g}$ with

$$\hat{g}(x) = \log \prod_{n=0}^{\infty} ((a \circ h^n)(x))^{-1}. \tag{3.2.25}$$

With (21) and (22) we have:

$$|\hat{g}(x)| \le \sum_{n=0}^{\infty} \left|\log(a \circ h^n)(x)\right| \le \sum_{n=0}^{\infty} C_1 \left|h^n(x)\right| \le C_1 |x| \sum_{n=0}^{\infty} q^n < \frac{C}{1-q}, \; x \in U_{\delta_4}$$

where

$$C = C_1 \sup_{x \in U_{\delta_4}} |x|.$$

Thus the series

$$\hat{g}(x) = \sum_{n=0}^{\infty} \log(a \circ h^n)(x)$$

converges independently from $x \in U_{\delta_4}$, i.e. the convergence is locally uniform. Thus $\hat{g}$ and using (24) and (25)

$$g(x) = g_0\, e^{\hat{g}(x)}, \; x \in U_{\delta_4}, \; g_0 \in \mathbf{R^1}$$

are real-analytic in a neighbourhood of O. From (24) it follows that the product (23) represents a nontrivial solution of (3.1.15). For $\mu = 0$ the assertion of the lemma is thus shown.

Let $\mu = 1, 2, 3, \ldots$. As h and a are assumed to be real-analytic $\exists\, \delta > 0$ such that the functional equation

$$h(x)^{\mu}\, (g_l \circ h)(x) = x^{\mu}\, a(x)\, g_l(x), \; x \in U_{\delta}$$

for the unknown function $g_l(x)$ can be considered. As $h_0 > 0$ by hypothesis $\exists\, \delta_1 > 0, \delta > \delta_1 > 0$, with $|h(x)| > 0$, $x \in U_{\delta_1}$, i.e. there follows

$$(g_1 \circ h)(x) = \left(\frac{x}{h(x)}\right)^{\mu} a(x) g_1(x),\ x \in U_{\delta_1}. \tag{3.2.26}$$

With

$$a^*(x) = \left(\frac{x}{h(x)}\right)^{\mu} a(x)$$

the solution g_1 of (26) and the solution g of (3.1.15) are related by

$$g_1(x) = x^{-\mu} g(x). \tag{3.2.27}$$

With $a(x) = a^*(x)$ in (17) we find:

$$\mu = \frac{\log a^*(0)}{\log h_0} = 0.$$

Using the proof for $\mu = 0$ we conclude that there exists a neighbourhood U, in which (26) has a nontrivial one-parametric real-analytic solution $g_1(x)$, $x \in U$. With (27) it follows that (3.1.15) has a nontrivial real-analytic solution g in a neighbourhood of O with a zero of order μ.

◊

Proof of Theorem 3.5: In Lemma 3.2 Theorem 3.5 has been proved under the additional assumption (18). As $h_0 \neq 1$ is excluded by hypothesis, it remains to consider $h_0 > 1$. As shown in the proof of Theorem 3.4, there exists U_δ, $\delta > 0$, in which $(a \circ h^{-1})(x)$ does not vanish and in which the real-analytic function $\tilde{h}(x) = h^{-1}(x)$, $x \in U_\delta$ exists. With $f(x) = g(x)$ and $b(x) = 0$ in (16) we have:

$$(g \circ \tilde{h})(x) - \tilde{a}(x) \cdot g(x) = 0,\ x \in U_\delta \tag{3.2.28}$$

where

$$\tilde{a}(x) = \frac{1}{(a \circ \tilde{h})(x)}.$$

$\tilde{h}$ and $\tilde{a}$ are real-analytic and, using (14) und (15), it is concluded that the assertions of Lemma 3.2 hold for (28). As in the proof of Theorem 3.4, the transition from (28) to (3.1.15) is considered. Thus, the assertions of Theorem

3.5 are also satisfied for (3.1.15).

◊

Theorem 3.6: Let the coefficient function h, a, b of (2.3.2) be real-analytic in O represented by the series (3.2.1). It is assumed that $h_0 > 0$, $h_0 \neq 1$ and $a_0 > 0$. Furthermore let

$$\mu = \frac{\log a_0}{\log h_0} \in \mathbf{N_0}.$$

Then, (2.3.2) has a one-parametric family of solution in the origin and we have

$$f(x) = O(x^{\mu} \log|x|),\ x \in U_\delta. \qquad (3.2.29)$$

Proof: The assumptions of the Theorem 3.5 are satisfied, (i.e. the corresponding homogeneous problem (3.1.15) of (2.3.2) has a one-parametric family of non trivial real-analytic solution g with a zero of order μ in O). Furthermore, using Theorem 3.3, there exists a formal power series

$$E(x) = \sum_{n=0}^{\infty} e_n x^n$$

that satisfies (3.1.24), and with Corollary 3.2 it follows that E(x) is real-analytic in x = O. The transition from (3.1.24) to (3.1.23) shows that the function f defined by (3.1.23) is a solution of (2.3.2) which satisfies (29) in a neighbourhood of the origin O.

◊

Concluding remarks: The discussion of (2.3.2) with real-analytic coefficient functions is concluded. The theorems describe the solution f of (2.3.2) in a neighbourhood of the fixed point h(O) = O and with the consideration (3.1.3) at the beginning of Section 3.1 also in a neighbourhood of a fixed point h(s) = s, $s \in \mathbf{R}^1$. The results can be summarized as follows:

1. $\mu \neq 0, 1, 2, \ldots$: Under the assumptions of Theorem 3.4 there exists a unique real-analytic solution in s.

2. $\mu = 0, 1, 2, \ldots.$: Under the assumptions of Theorem 3.6 there exists no real-analytic solutions in s. For all solution in s the representation (29) holds.

4. The regularity of the 2π-periodic solution

4.1. The continuation of a continuous function by the linear functional equation

Let $[s, t]$, $s, t \in \mathbf{R}^1$, $s < t$ be a closed interval on the real axis and $C[s, t]$ denotes the set of the continuous, real-valued functions defined on $[s, t]$. We consider

$$f(h(x)) - a(x) \cdot f(x) = b(x) \qquad (4.1.1)$$

for all $x \in [s, t]$ with $a, b, h \in C[s, t]$ for the unknown function f. Furthermore, let h be invertible on $[s, t]$ with fixed points in s and t, i.e. we have $h^n(x) \in [s, t]$, $x \in [s, t]$, $n = 0, 1, 2,\ldots$. We derive theorems that are used in Section 4.2 for the detailed discussion of the unique, continuous, 2π-periodic solution f described in Theorem 2.1. In this section, we only consider continuous coefficient functions h, a and b without any 2π-periodicity assumptions.

Theorem 4.1: Let the following assumptions hold for (1):
1. $h, a, b \in C[s, t]$, $t > s$;
2. $h(s) = s$, $h(x) < x$ for $x \in (s, t)$, $h(t) = t$, h invertible on $[s, t]$;
3. For $x_0 \in (s, t)$ let $f_1 \in C[h(x_0), x_0]$ be a function that satisfies (1) in $x = x_0$:

$$f_1(h(x_0)) = a(x_0) \cdot f_1(x_0) + b(x_0).$$

Then f_1 can be uniquely extended to $[h^n(x_0), x_0]$, $n = 1, 2, 3,\ldots$ by (1), i.e. there exists a unique function $f_n \in C[h^n(x_0), x_0]$ such that

$$\begin{aligned} &1.\ f_n \text{ satisfies (1), } \forall\, x \in [h^{n-1}(x_0), x_0],\ n = 1, 2, 3,\ldots; \\ &2.\ f_n(x) = f_1(x),\ \forall\, x \in [h(x_0), x_0]. \end{aligned} \qquad (4.1.2)$$

Proof: Let $x_n = h^n(x_0)$. We proof the theorem by induction with respect to n. For $n = 1$ the assertion follows from assumption 3. Let $f_n \in C[x_n, x_0]$, $n = 1, 2,\ldots$ with (1) be given. As h is invertible any continuation is of the form

$$f_{n+1}(x) = \begin{cases} f_n(x), x \in [x_n, x_0] \\ \\ a(h^{-1}(x)) \cdot f_n(h^{-1}(x)) + b(h^{-1}(x)), x \in [x_{n+1}, x_n) \end{cases}, \qquad (4.1.3)$$

i.e. the continuation is unique. In the following we show $f_{n+1} \in C[x_{n+1}, x_0]$. For $x \in [x_{n+1}, x_n)$ and $x \in (x_n, x_0]$ the continuity follows from (3), assumption 1 and $f_n \in C[x_n, x_0]$. It remains to demonstrate the continuity in x_n. From (3) we find

$$f_{n+1}(x) = f_n(x),\ x \in [x_n, x_0].$$

From $f_n \in C[x_n, x_0]$, we have:

$$f_{n+1} \text{ is continuous in } x_n \text{ as x approaches from the right.} \qquad (4.1.4)$$

We consider the limit from the left

$$\lim_{y \to x_n^-} f_{n+1}(y)$$

and, using assumption 2, let $x = h^{-1}(y)$ with $y \in [x_{n+1}, x_n)$. The continuity of h yields

$$\lim_{y \to x_n^-} f_{n+1}(y) = \lim_{x \to x_{n-1}^-} f_{n+1}(h(x)).$$

Because of $y = h(x) \in [x_{n+1}, x_n)$, (3) and assumption 2 it follows

$$f_{n+1}(h(x)) = a(x) \cdot f_n(x) + b(x).$$

As $f_n \in C[x_n, x_0]$ satisfies (1) by induction assumption we conclude with assumption 1

$$\lim_{x \to x_{n-1}^-} f_{n+1}(h(x)) = a(x_{n-1}) \cdot f_n(x_{n-1}) + b(x_{n-1}) = f_n(x_n).$$

$x = x_n$ in (3) gives:

$$f_{n+1}(x_n) = f_n(x_n). \qquad (4.1.5)$$

f_{n+1} is thus in x_n continuous from left and, using (4), we find that f_{n+1} is continuous in x_n. For $x \in (x_{n+1}, x_0]$ condition (2) follow from the induction assumption and (3). For $x = x_{n+1}$ we have from (3) and (5):

$$f_{n+1}(x_{n+1}) = a(x_n) \cdot f_n(x_n) + b(x_n) = a(x_n) \cdot f_{n+1}(x_n) + b(x_n),$$

i.e. condition 1 in (2) is fulfilled in x_{n+1}.

◊

Lemma 4.1: Let the following assumptions hold for (1):

1. $h, a, b \in C[s, t], t > s$;
2. $h(s) = s, h(x) < x$ for $x \in (s, t), h(t) = t$, h invertible on $[s, t]$;
3. f_n is the continuous continuation of a continuous function f_l on the closed interval $[h^n(x_0), x_0]$, $n = 1, 2, \ldots$ according to Theorem 4.1;
4. $A = \sup\limits_{x \in [s, x_0]} |a(x)|$.

Then we have

$$\left|f_{n+1}(h^n(x))\right| \le A^n \left|f_1(x)\right| + B_x(1 + A + \ldots + A^{n-1}), \qquad x \in [x_1, x_0], n = 0, 1, 2, \ldots, \tag{4.1.6}$$

where

$$B_x = \sup_{y \in [s, x]} |b(y)|, \; x \in [s, x_0]. \tag{4.1.7}$$

Proof: We prove the lemma by induction on n. For n = 0 the assertion is trivial. Let $f_n \in C[x_n, x_0]$ with (6) be given. From $h(x) \le x$, $x \in [s, x_0]$ it follows that

$$B_{h^n(x)} \le B_x, \; n = 0, 1, 2, \ldots, \; x \in [x_1, x_0]. \tag{4.1.8}$$

As f_{n+1} satisfies (1) on $[x_n, x_0]$, we have with (3):

$$f_{n+1}(h^n(x)) = a(h^{n-1}(x)) \cdot f_{n+1}(h^{n-1}(x)) + b(h^{n-1}(x)) =$$

$$a(h^{n-1}(x)) \cdot f_n(h^{n-1}(x)) + b(h^{n-1}(x)), \; x \in [x_1, x_0].$$

By using (6), (7) and (8), we obtain:

$$\left|f_{n+1}(h^n(x))\right| = \left|a(h^{n-1}(x)) \cdot f_n(h^{n-1}(x)) + b(h^{n-1}(x))\right| \le$$

$$A\left(A^{n-1}\left|f_1(x)\right| + B_x(1 + A + \ldots + A^{n-2})\right) + B_{h^{n-1}(x)} \le$$

$$A^n\left|f_1(x)\right| + B_x(1 + A + \ldots + A^{n-1}).$$

◊

Theorem 4.2: Let the following assumptions hold for (1):
1. $h, a, b \in C[s, t], t > s$;
2. $h(s) = s$, $h(x) < x$ for $x \in (s, t)$, $h(t) = t$, h invertible on $[s, t]$;
3. $f_l(x) \in C[h(x_0), x_0]$, $x_0 \in (s, t)$, satisfies (1) in x_0:

$$f_l(h(x_0)) = a(x_0) \cdot f_l(x_0) + b(x_0);$$

4. $\left|a(x)\right| < 1$, $x \in [s, x_0]$.

Then f_l can be uniquely extended to a continuous solution f of (1) that is defined on $[s, x_0]$.

Proof: Let us first assume

$$b(s) = 0. \tag{4.1.9}$$

Equation (7) and the assumption $b \in C[s, t]$ yield

$$\lim_{x \to s^+} B_x = 0.$$

Let $\varepsilon > 0$ be given. It follows that $\exists\ \delta > 0$ and $0 < A < 1$ such that

$$B_x \le \frac{1}{2}(1 - A)\,\varepsilon,\ x \in [s, s + \delta] \tag{4.1.10}$$

and

$$\left|a(x)\right| < A < 1, x \in [s, x_0].$$

With $x_n = h^n(x_0)$ we have by assumption $x_n \to s$, i.e. there exists $n_l \in \mathbf{N}$ with $[x_{n_1}, x_{n_1-1}] \subset [s, s + \delta)$. Let $y_0 = x_{n_1-1}$ and $y_1 = h(y_0)$. We consider the unique continuous continuation $f_{n_1}(x)$ of $f_l(x)$ on $[y_1, x_l]$ described in Theorem 4.1 and g_l denotes the restriction of f_{n_1} to $[y_1, y_0]$. We choose $n_2 \in \mathbf{N}$ such that

$$A^n < \frac{\varepsilon}{2 \cdot \sup\limits_{x \in [y_1, y_0]} \left|g_1(x)\right|}, n \geq n_2.$$

We use Lemma 4.1 with $x_0 = y_0$ and $f_n(x) = g_n(x)$, n = 1, 2,….. From (6), (10) and assumption 4 of the theorem we obtain:

$$\left|g_{n+1}(h^n(x))\right| \leq A^n \left|g_1(x)\right| + B_x \frac{1-A^n}{1-A} < \varepsilon, n \geq n_2, x \in [y_1, y_0]$$

and it follows

$$\lim_{n\to\infty} g_{n+1}(h^n(x)) = 0, x \in [y_1, y_0].$$

As $x \in [y_1, y_0]$ can be considered as the $(n_l - 1)$-times iteration of an element from $[x_l, x_0]$, we have

$$\lim_{n\to\infty} g_{n+1}(h^n(h^{n_1-1}(x))) = 0, x \in [x_l, x_0].$$

As $g_l(x), x \in [y_1, y_0]$ and $f_{n_1}(x), x \in [y_1, y_0]$ are identical it follows that their continuous continuation are also identical. It is concluded that

$$f_{n+n_1}(h^n(h^{n_1-1}(x))) = g_{n+1}(h^n(h^{n_1-1}(x))), x \in [x_l, x_0], n = 0, 1, 2,…..$$

Thus

$$\lim_{n\to\infty} f_{n+1}(h^n(x)) = 0, x \in [x_1, x_0].$$

From (1), $h(s) = s$, $b(s) = 0$ and the assumption $a(s) \neq 1$ we find $f(s) = 0$, i.e. f_l can be continuously extended as x approaches from right. The theorem is therefore proved under the additional assumption (9). For $b \in C[s, t]$ we consider

$$d(x) = b(x) - \frac{1-a(x)}{1-a(s)} b(s) \qquad (4.1.11)$$

and find

$$d(s) = 0.$$

From a, b $\in$ C[s, t] and assumption 4 we have d $\in$ C[s, t]. From the first part of the proof it is concluded that

$$f^*(h(x)) = a(x) \cdot f^*(x) + d(x)$$

has a solution f* $\in$ C[s, x_0] of (1) and it is easily verified with (11) that the function, defined by

$$f(x) = f^*(x) + \frac{b(s)}{1-a(s)}, \; x \in [s, x_0]$$

is a solution of (1).

◊

Theorems 4.1 and 4.2 show that any given continuous function defined on the interval [h(x_0), x_0], x_0 $\in$ (s, t) and satisfying the functional equation (1) has a unique continuation on the interval [s, x_0]. These solutions are continuous in s and we have

$$f(s) = \frac{b(s)}{1-a(s)}.$$

The following theorem deals with the solutions of (1) on the closed interval [s, t].

Theorem 4.3: Let the following assumptions hold for (1):
1. h, a, b $\in$ C[s, t], t > s;
2. h(s) = s, h(x) < x for x $\in$ (s, t), h(t) = t, h invertible on [s, t];
3. $|a(t)| < 1$.
Then there exists a unique solution f on (s, t]. f is represented by (2.3.7).
We consider 2 cases:
1. If $|a(s)| < 1$ then f can be uniquely extended in the fixed point s, i.e. equa-

tion (1) has a unique solution $f \in C[s, t]$ and (2.3.7) yields

$$f(s) = \frac{b(s)}{1 - a(s)}.$$

2. If $|a(s)| > 1$ and $b(x) \neq 0$, $\forall x \in [s, t]$ then f is unbounded in s, i.e. equation (1) has no continuous solution on [s, t].

Proof: First we show assertion 1 of the theorem. The proof consists of two parts.

Step 1:

Because of $|a(s)| < 1$, $|a(t)| < 1$ and $a \in C[s, t]$ it follows that $\exists\ \delta > 0$ and $\exists\ A \in \mathbf{R}^1$ with

$$|a(x)| < A < 1,\ \forall x \in [s, s + \delta],\ \forall x \in [t - \delta, t]. \tag{4.1.12}$$

We show that the series (2.3.7) represents a unique solution $f \in C[t - \delta, t]$ of (1). Iteration of h and assumption 2 of the theorem yield

$$h^{-n-1}(x) \geq h^{-n}(x),\ n = 0, 1, 2, \ldots,\ \forall x \in [t - \delta, t]. \tag{4.1.13}$$

Thus, the functions a and b in the partial sums of (2.3.7) are only evaluated for $x \in [t - \delta, t]$. As in (2.3.18) we find with (12)

$$|f(x)| \leq M\,(1 + A + A^2 + \ldots) = \frac{M}{1 - A} < \infty,$$

where

$$M = \sup_{x \in [t-\delta, t]} |b(x)|.$$

The partial sums of (2.3.7) converge uniform, (i.e. $f \in C[t - \delta, t]$). To prove uniqueness of f, let $g \in C[t - \delta, t]$ be a second solution of (1) defined on $[t - \delta, t]$. Because of (13) we consider (2.3.19) for $x \in [t - \delta, t]$ and (2.3.15) follows from (12) by letting $n \to \infty$. It is concluded that there exists a unique solution $f \in C[t - \delta, t]$.

Step2:

Let $x_0 = h^{-1}(t - \delta)$ and $x_n = h^n(x_0)$. As s is an attractive fixed point $\exists\ n_l \in \mathbf{N}$ with

$$[x_{n_1}, x_{n_1-1}] \subset [s, s+\delta). \tag{4.1.14}$$

The restriction of the unique solution $f \in C[t - \delta, t]$ derived in step 1 to $[x_1, x_0]$ is denoted with f_l. Using Theorem 4.1 let f_{n_1} be the unique continuous continuation of f_l to $[h(x_{n_1-1}), x_0]$. Based on (3) induction on n yields:

$$f_n(x_n) = \prod_{k=0}^{n-1} a(x_k) f_1(x_0) + \sum_{k=0}^{n-1} \prod_{j=k+1}^{n-1} a(x_j) b(x_k), \; n = 1, 2, \ldots.. \tag{4.1.15}$$

Let g_l be the restriction of f_{n_1} to $[h(y_0), y_0]$ where $y_0 = x_{n_1-1}$. From (12) and (14) we conclude with $x_0 = y_0$, $f_l(x) = g_l(x)$ and (12) in Theorem 4.2 that there exists a unique continuous continuation of g_l to $[s, y_0]$. The continuation of g_l is also a unique continuation of f_n given by (15) and as a consequence also of the unique solution $f \in C[t - \delta, t]$ in the neighbourhood of s. By Theorem 4.2 it is concluded that $f \in C[t - \delta, t]$ can be extended to a unique solution $f \in C[s, t]$ of (1). Furthermore (2.3.16) shows that f satisfies (2.3.7) formally and consequently *the series (2.3.7) represents a unique solution of (1) on [s, t]*. The assertion of case 1 is thus proved and it remains to demonstrate case 2.

Because of $|a(s)| > 1$, $|a(t)| < 1$ and $a \in C[s, t]$ it follows that $\exists\ \delta > 0$ and $\exists\ A > 0, B > 0$ with

$$|a(x)| > 1, x \in [s, s+\delta], \; |a(x)| < A < 1, x \in [t-\delta, t] \tag{4.1.16}$$

and

$$|b(x)| > B, \; \forall x \in [s, t].$$

The existence of a unique solution $f \in C[t - \delta, t]$, $\delta > 0$ of (1) that satisfy (2.3.7) has already been shown in step 1 of the proof in case 1.

As s is an attractive fixed point, it follows $\exists\ n_l \in \mathbf{N}$ with (14) such that (16) holds. Let $y_1 = x_{n_1}$, $y_0 = x_{n_1-1}$. Using Theorem 4.1 let f_{n_1} denote the unique continuation on $[y_1, x_1]$ given by (15). As the continuations f_n, $n = 1, 2, \ldots, n_l$ satisfies (1) by construction, it is seen with (2.3.16) that f_n is represented by the series (2.3.7). It is concluded with (16) that $f \in C[t-\delta, t]$ can be extended in a neighbourhood of s that excludes s to a solution of (1) that is represented by (2.3.7).

Let g_l be the restriction of f_{n_1} to $[y_1, y_0]$. Using again Theorem 4.1, let g_n, n = 1, 2,…. be the continuous continuation of g_l on the interval $[y_n, y_0]$. With $f_n(x) = g_n(x)$ and $x_n = y_n$ in (15) we find:

$$g_n(y_n) = \prod_{k=0}^{n-1} a(y_k) g_1(y_0) + \sum_{k=0}^{n-1} \prod_{j=k+1}^{n-1} a(y_j) b(y_k), \; n = 1, 2, \ldots. \qquad (4.1.17)$$

We have to consider the following cases:

1. $a(x) > 1$, $\forall x \in [s, s+\delta]$

We distinguish three cases:

A. $b(x) > B$, $\forall x \in [s, t]$, $g_l(y_0) \geq 0$

From $y_0 \in [s, s+\delta]$ it follows by neglecting the first part of the sum in (17):

$$g_n(y_n) > \sum_{k=0}^{n-1} b(y_k) > n \cdot B, \; n = 0, 1, 2, \ldots.. \qquad (4.1.18)$$

B. $b(x) < -B$, $\forall x \in [s, t]$, $g_l(y_0) \leq 0$

As in case A we have:

$$g_n(y_n) < \sum_{k=0}^{n-1} b(y_k) < -n \cdot B, \; n = 0, 1, 2, \ldots.. \qquad (4.1.19)$$

From (18) and (19) it follows that the unique extension of g_n of g_l and also the unique extension f_n of f is unbounded on $[s, s+\delta]$. As $f \in C[t-\delta, t]$ is unique on $[t-\delta, t]$ and a continuous function is necessarily bounded it is concluded that there exists no continuous solution of (1) on [s, t].

Let the *signum function* for a real number $c \in \mathbf{R}^1$ be defined by

$$\operatorname{sign} c = \begin{cases} 1, & c > 0 \\ 0, & c = 0\,. \\ -1, & c < 0 \end{cases}$$

It remains to discuss the case

iC. There is a x such that sign $b(x) \neq$ sign $g_l(y_0)$

We consider the unique solution $f \in C[t - \delta, t]$, $\delta > 0$. With $|a(t)| < 1$ and

$$f(t) = \frac{b(t)}{1 - a(t)}$$

it follows that $\exists\ \delta_1 > 0$ with

$$\text{sign } f(x) = \text{sign } b(x),\ x \in [t - \delta_1, t].$$

As shown in Theorem 4.1 the continuation $f_{n_1}(x)$, $x \in [x_{n_1}, x_0]$ of $f(x)$, $x \in [t - \delta_1, x_0]$ is continuous and unique. As $f_{n_1}(x_{n_1}) = g_l(x_0)$ it follows by hypothesis that sign $f_{n_1}(x_{n_1})$ = sign $b(x)$ cannot hold on $[x_{n_1}, x_0]$ and consequently $\exists\ \tilde{x} \in [s, t]$ with

$$f_{n_1}(\tilde{x}) = 0.$$

Using (1) we have

$$f_{n_1}(h(\tilde{x})) = b(\tilde{x}). \tag{4.1.20}$$

As $b(x) \neq 0$, $\forall x \in [s, t]$ by assumption of the theorem we find with $x_0 = h(\tilde{x})$

$$f_{n_1}(x_0) \neq 0$$

and

$$\text{sign } f_{n_1}(x_0) = \text{sign } b(\tilde{x}).$$

We consider (17) with $x_0 = y_0$ and conclude that for the continuation starting in $f(x_0) = g_l(y_0)$ either the previously discussed cases A or B apply and from either (18) or (19) follows the assertion of the theorem.

2. $a(x) < -1$, $\forall x \in [s, s + \delta]$

We again consider the continuous continuation of the unique solution $f \in C[t - \delta, t]$, $\delta > 0$ on the interval (14) and its restriction g_l to $[y_l, y_0]$ where $y_l = x_{n_1}$, $y_0 = x_{n_1\text{-}1}$. We consider the following cases:

A. $b(x) > B,\ \forall x \in [s, t],\ g_l(y_0) > 0$

From $a(x) < -1,\ x \in [s, s + \delta]$ it follows that the partial sums of (20) have alternating signs, i.e. the following estimations holds:

$$g_{2n-1}(y_{2n-1}) > \sum_{k=0}^{n} b(y_{2k}) > (2n-1)\cdot B,\ n = 1, 2,\ldots. \tag{4.1.21a}$$

and

$$g_{2n}(y_{2n}) < \sum_{k=0}^{n} b(y_{2k+1}) < -2n\cdot B,\ n = 1, 2,\ldots.. \tag{4.1.21b}$$

B. $b(x) < -B,\ \forall x \in [s, t],\ g_l(y_0) < 0$

Analogous to (21) we find

$$g_{2n-1}(y_{2n-1}) < \sum_{k=0}^{n} b(y_{2k}) < (2n-1)\cdot B,\ n = 1, 2,\ldots. \tag{4.1.22a}$$

and

$$g_{2n}(y_{2n-1}) > \sum_{k=0}^{n} b(y_{2k+1}) > -2n\cdot B,\ n = 1, 2,\ldots.. \tag{4.1.22b}$$

C. $\operatorname{sign} b(x) \neq \operatorname{sign} g_l(y_0),\ \forall x \in [s, t]$

Following (20) let $\tilde{y}_0 = y_0$ with $f_{n_1}(\tilde{x}) = g_l(\tilde{y}_0) = 0$. Applying the sequence $\tilde{y}_n = h^n(y_0)$ to (17) yields

$$g_n(\tilde{y}_n) = \sum_{k=0}^{n-1} \prod_{j=k+1}^{n-1} a(\tilde{y}_j) b(\tilde{y}_k),\ n = 1, 2,\ldots,$$

i.e. either (21) or (22) apply. For $a(s) < -1$ it is concluded that the solutions of (1) are *oscillating and unbounded* in the neighbourhood of s. The proof of Theorem 4.3 is thus completed.

◊

From the first part of proof of Theorem 4.3 it follows that in a left side neighbourhood of t there exists a unique continuous solution f of (1). Under the assumptions of Theorem 3.4, this solution can be extended to a real-analytic solution defined in a neighbourhood of t.

Theorem 4.2 shows that (1) has many continuous solutions on the interval $[s, t)$. However, on the closed interval $[s, t]$ either there exists a unique solution f or there is no continuous solution (Theorem 4.3). By using a symmetry argument the assertion of Theorem 4.3 can be proven for functions h with a repelling fixed point s and an attractive fixed point t.

Theorem 4.4: Let the following assumptions hold for (1):
1. $h, a, b \in C[s, t], t > s$;
2. $h \in C[s, t], h(s) = s, h(x) > x$ for $x \in (s, t), h(t) = t$, h invertible on $[s, t]$;
3. $|a(s)| < 1$.

Then, there exists a unique solution f on $[s, t)$ and the two cases considered in Theorem 4.3 apply to the fixed point t.

Proof: Starting point is again (1). With $t_1 = -t$, $s_l = -s$ and

$$\begin{aligned} h_1(x) &= -h(-x), \\ a_1(x) &= a(-x), \\ b_1(x) &= b(-x) \end{aligned} \tag{4.1.23}$$

for $x \in [t_1, s_l]$ we consider

$$(f_l \circ h_1)(x) = a_1(x) \cdot f_1(x) + b_1(x), \ x \in [t_1, s_l] \tag{4.1.24}$$

for the unknown function f_l. With $s = t_1$, $t = s_l$, $h(x) = h_1(x)$, $a(x) = a_1(x)$, $b(x) = b_1(x)$, $x \in [s, t]$ the coefficient functions of (24) satisfy the assumptions of Theorem 4.3, (i.e. (24) has a unique solution $f_l \in C[s, t]$ in case 1 or no continuous solution on $[s, t]$ in case 2). Let

$$f(x) = f_1(-x), \ x \in [s, t]. \tag{4.1.25}$$

This is the same as

$$f_1(x) = f(-x), \ x \in [t_1, s_l]. \tag{4.1.26}$$

From (23), (24) and (26) it follows

$$f_l(-h(-x)) = a(-x) \cdot f(-x) + b(-x),\ x \in [t_l, s_l].$$

With (25) we have

$$f(h(x)) = a(x) \cdot f(x) + b(x),\ x \in [s, t], \tag{4.1.27}$$

i.e. $f(x)$, $x \in [s, t]$ satisfies (1), and from the definition of f in (25), it follows that f satisfies the assertions of Theorem 4.3.

◊

Theorem 4.5: Let the following assumptions hold for (1):
1. $h, a, b \in C[s, t]$, $t > s$;
2. $h(s) = s$, $h(x) > x$ for $x \in (s, t)$, $h(t) = t$, h invertible on $[s, t]$;
3. $|a(t)| > 1$.
Then there exists a unique continuous solution f on $(s, t]$. f is represented by (2.3.6).

We consider 2 cases:
1. If $|a(s)| > 1$ then f can be uniquely extended in the fixed point s, i.e. equation (1) has a unique solution $f \in C\,[s, t]$ and (2.3.6) yields

$$f(s) = \frac{b(s)}{a(s) - 1}.$$

2. If $|a(s)| < 1$ and $b(x) \neq 0$, $x \in [s, t]$ then f is unbounded in s, i.e. equation (1) has no continuous solution on $[s, t]$.

Proof: Let us first consider case 1. From $|a(s)| > 1$, $|a(t)| > 1$ and $a \in C[s, t]$ it follows that $\exists\ \delta > 0$ and $\exists\ A \in \mathbf{R}^1$ with

$$|a(x)| > A > 1,\ x \in [s, s + \delta],\ x \in [t - \delta, t]. \tag{4.1.28}$$

Following the proof of step 1 in Theorem 4.3 it is concluded that there exists a unique solution of $f \in [t - \delta, t]$. f is represented by the series (2.3.6).

Let $\tilde{h}(x) = h^{-1}(x)$, $x \in [s, t]$ denote the inverse of h for $x \in [s, t]$ and let $x_0 = \tilde{h}(t - \delta)$. Furthermore, the sequence x_n, $n = 0, 1, 2, \ldots$ is given by $x_n = \tilde{h}^n(x_0)$.

As s is a repulsive fixed point $\exists\ n_l \in \mathbf{N}$ with $[x_{n_1}, x_{n_1 - 1}] \subset [s, s + \delta)$.

Replacing h by $\tilde{h}$ in Theorem 4.1 it follows that f can be uniquely continually extended to $[\tilde{h}(y_0), x_0]$ where $y_0 = x_{n_1-1}$. As the continuation satisfy (1) it follows with (2.3.8) that they satisfy (2.3.6). Based on (28), the functional equation

$$f(\tilde{h}(x)) - \tilde{a}(x) \cdot f(x) = \tilde{b}(x) \tag{4.1.29}$$

where

$$\tilde{a}(x) = \frac{1}{a(\tilde{h}(x))}, \tilde{b}(x) = -\frac{b(\tilde{h}(x))}{a(\tilde{h}(x))}, x \in [s, s+\delta], x \in [t-\delta, t]$$

can be considered. Let g_l be the restriction of f to $[\tilde{h}(y_0), y_0]$. With (28) and as $\tilde{a}, \tilde{b} \in C[s, s+\delta]$ it is concluded with Theorem 4.2 in case 1 that there exists a unique continuation of g_l and as a consequence also of $f \in C[t-\delta, t]$ in the fixed point s. It follows that there exists a unique solution $f \in C[s, t]$ of (29) and consequently of (1). As verified in (2.3.8) f is represented by the series (2.3.6). Case 1 of the theorem is thus demonstrated and we consider case 2.

As s is an attractive fixed point of $\tilde{h}$, it is seen with $|a(s)| < 1$ that $\exists \delta_1 > 0$ and $\exists A \in \mathbf{R}^1$ such that

$$|a(\tilde{h}(x))| < A < 1, x \in [s, s+\delta_1]. \tag{4.1.30}$$

Furthermore from $|b(\tilde{h}(x))| \neq 0$, $x \in [s, t]$ we conclude that $\exists B \in \mathbf{R}^1$ with $B > 0$

$$|b(\tilde{h}(x))| > B. \tag{4.1.31}$$

With $x_n = \tilde{h}^n(x_0)$, $n = 1, 2, \ldots.$ and $x_0 = t - \delta$ it follows $\exists n_l \in \mathbf{N}$ with

$$[x_{n_1}, x_{n_1-1}] \subset [s, s+\delta_1]. \tag{4.1.32}$$

As shown in the proof of case 1 there exists a unique solution $f \in C[t-\delta, t]$ that can be uniquely extended to the interval $[x_{n_1}, x_{n_1-1}]$ with (32). We consider (29) for $x \in [s, s+\delta_1]$. Combining (30) and (31) yields

$$|\tilde{a}(s)| > \frac{1}{A} > 1,$$

$$\left|\tilde{b}(s)\right| > \frac{B}{A} > 0,\ x \in [s, s + \delta_1],\ \text{respectively}$$

for the coefficient functions $\tilde{a}, \tilde{b}$ in (29) and, as $\tilde{a}, \tilde{b} \in C[s, s + \delta]$, it follows that (29) satisfies in the fixed point s the assumptions of Theorem 4.3 in case 2. The application of the discussion in the proof of case 2 in Theorem 4.3 shows that the continuous continuation of the unique solution f in the fixed point t is unbounded in the neighbourhood of s, i.e. there exists no continuous solution of (29) on the interval [s, t]. The transition from (29) to (1) completes the proof of the theorem.

◊

Following the proof of Theorem 4.4 by applying Theorem 4.5 yields the following corollary:

Corollary 4.1: Let the following assumptions hold for (1):
1. $h, a, b \in C[s, t]$, $t > s$;
2. $h(s) = s$, $h(x) < x$ for $x \in (s, t)$, $h(t) = t$, h invertible on [s, t];
3. $|a(s)| > 1$.

Then, there exists a unique continuous solution f on [s, t) and the two cases considered in Theorem 4.5 apply to the fixed point t.

◊

4.2. The case with fixed points

In this section the coefficient functions h, a and b of (4.1.1) are assumed to be defined on the real axis, i.e. we consider

$$(f \circ h)(x) - a(x) \cdot f(x) = b(x) \tag{4.2.1}$$

for all $x \in \mathbf{R}^1$. Under the assumptions of Theorem 2.1 the linear functional equation (1) has a unique solution $f \in P$. In Section 4.2 and 4.3 the regularity of f is discussed in detail. This section deals with the case that *h has fixed points.* More precisely, the following is assumed for the coefficient functions h and a in (1).

Assumption 4.1: Let a and h in (1) satisfy the following:

1. $a \in P$ and $h(x) = x + p(x)$, $p \in P$ is invertible from $\mathbf{R}^1 \to \mathbf{R}^1$.

2. h has N fixed points $s_0, \ldots, s_{N-1}$, $N \geq 2$, N even, with $s_n \in [0, 2\pi)$, $0 \leq n \leq N-1$ and

$$\left| a(s_n) \right| \neq 1. \tag{4.2.2}$$

3. h is differentiable in the fixed points s_n, $0 \leq n \leq N-1$ with $h'(s_n) > 0$, $h'(s_n) \neq 1$ and the derivatives in s_n, $0 \leq n \leq N-1$ are related by

$$\log h'(s_{n-1}) \cdot \log h'(s_n) < 0,\ 1 \leq n \leq N-1.$$

From condition 2 in the above assumption follows for $k \in \mathbf{Z}$ and $0 \leq n \leq N-1$:

$$h(s_n + 2k\pi) = s_n + 2k\pi + p(s_n + 2k\pi) = s_n + 2k\pi + p(s_n) = s_n + 2k\pi,$$

i.e. h has fixed points in

$$s_n + 2k\pi,\ 0 \leq n \leq N-1,\ k \in \mathbf{Z} \tag{4.2.3}$$

and furthermore using (2) we have

$$\left|a(s_n + 2k\pi)\right| \neq 1,\ 0 \le n \le N-1,\ k \in \mathbf{Z}.$$

Definition 4.1: Let a and h satisfy Assumption 4.1. The *characteristic exponent* μ_{s_n} in the fixed point s_n of h is defined by

$$\mu_{s_n} = \frac{\log\left|a(s_n)\right|}{\log h'(s_n)},\ 0 \le n \le N-1. \tag{4.2.4}$$

Remarks:

1. Using Assumption 4.1 and (3) we have $\mu_{s_n + 2k\pi} = \mu_{s_n}$, $k \in \mathbf{Z}$, $0 \le n \le N-1$.
2. In Sections 3.1 and 3.2 it is assumed that h(O) = O. Assuming a(O) > 0 in (3.1.13) and (3.2.17), the characteristic exponent of h in the fixed point O are introduced.
3. The characteristic exponents are the log ratio of the two eigenvalues of the Jacobian in the fixed point s_n. This can be shown by considering (2.2.2b) and (2.2.2c) followed by solving the first derivative of the condition of invariance for the derivative of the invariant curve in the fixed point s_n (see also Aronson, [2, page 334]).

In the following the characteristic exponents are used for the investigation of the unique 2π-periodic solution f described in Theorem 2.1. We consider *the case with fixed points* introduced in Assumption 4.1 throughout this section.

Theorem 4.6: It is assumed that h and a in (1) satisfy Assumption 4.1. Furthermore let $b \in P$.

Then the following holds:

1. If

$$\mu_{s_{n-1}} \cdot \mu_{s_n} < 0,\ 1 \le n \le N-1 \tag{4.2.5}$$

then (1) has a unique solution $f \in P$.

2. If $\exists n_0$, $0 \le n_0 \le N-1$ such that

$$\mu_{s_{n_0}} < 0,\ \mu_{s_{n_0+1}} < 0 \tag{4.2.6a}$$

and

$$b(x) \neq 0, \ \forall x \in [s_{n_0}, s_{n_0+1}] \qquad (4.2.6b)$$

then (1) has no continuous solution on $[0, 2\pi]$.

First we prove the following lemma.

Lemma 4.2: Let the assumption of Theorem 4.6 be satisfied. We consider two cases:

1. The assertion of Theorem 4.6 in case 1 follows under the additional assumption

$$|a(s_0)| < 1. \qquad (4.2.7)$$

2. If

$$|a(s_{n_0+1})| < 1 \qquad (4.2.8)$$

for the index n_0, $0 \le n_0 \le N-1$ in (6), the assertion in case 2 of Theorem 4.6 holds.

Proof: As it is assumed that h^{-1} exists on $\mathbf{R}^1$ and h satisfies Assumption 4.1, the inverse of h exists also on the interval $[s_n, s_{n+1}]$, $0 \le n \le N-1$. In addition we have either

$$h(x) < x \qquad (4.2.9)$$

or

$$h(x) > x, \ x \in (s_{n-1}, s_n), \ 1 \le n \le N. \qquad (4.2.10)$$

From $h(x) = x + p(x)$, $p \in P$, $a \in P$, $b \in P$ it follows

$$h, a, b \in C[s_{n-1}, s_n], \ 1 \le n \le N. \qquad (4.2.11)$$

Proof of Lemma 4.2 in case 1: From condition 3 in Assumption 4.1 and, using condition (5) for $n = 1$ with (7), we have:

$$\log|a(s_1)| < 0,$$

i.e.

$$|a(s_1)| < 1.$$

Applying condition 3 of Assumption 4.1 to the fixed points $s_2,\ldots,s_n$ yields with (5) and (7):

$$|a(s_n)| < 1,\ 1 \le n \le N. \tag{4.2.12}$$

With $s = s_{n-1}$, $t = s_n$, $1 \le n \le N$ in Theorem 4.3, Theorem 4.4, respectively in case 1 and, using (9), (10) and (11), it follows that on the interval $[s_{n-1}, s_n]$, $1 \le n \le N$ the assumptions of Theorem 4.3, Theorem 4.4, respectively are satisfied. It is concluded that there exists a unique solution $f \in [s_{n-1}, s_n]$ of (1). In addition (2) and (3) with k = 1 yield

$$f(s_n) = \frac{b(s_n)}{1 - a(s_n)},\ 0 \le n \le N,$$

i.e. by using (2) and $a \in P$, $b \in P$, there exists a unique solution $f(x)$, $x \in [s_0, s_N]$ of (1) with

$$f(s_0) = f(s_0 + 2\pi) = f(s_N)$$

and, as the continuous 2π-periodic extension of f to $\mathbf{R}^1$ is unique, the assertion $f \in P$ is thus demonstrated.

Proof of Lemma 4.2 in case 2: Using 3 in Assumption 4.1 we find with (6a) and (8)

$$h'(s_{n_0}) < 1,\ h'(s_{n_0+1}) > 1 \tag{4.2.13}$$

for the index n_0. With (6a) and (4) it is concluded $|a(s_{n_0})| > 1$, i.e. with $s = s_n$, $t = s_{n+1}$, (6), (8), (11) and (13) the assumptions of Theorem 4.3 in case 2 are satisfied. It follows that (1) has no continuous solution on $[s_{n_0}, s_{n_0+1}]$ and consequently there exists no continuous solution of (1) defined on $\mathbf{R}^1$. Lemma 4.2 is thus demonstrated.

◊

Proof of Theorem 4.6 in case 1: From (5) we have $\mu_{s_0} \neq 0$, thus $|a(s_0)| \neq 1$ and, as the assertion has already be shown under the condition (7), it remains to consider

$$|a(s_0)| > 1$$

and as in (12) it is concluded

$$|a(s_n)| > 1, 0 \leq n \leq N.$$

By using Theorem 4.5 and Corollary 4.1, the assertion follows as in the second part of the proof of Lemma 4.2 in case 1.

Proof of Theorem 4.6 in case 2: From $\mu_{s_{n_0+1}} < 0$ for the index n_0 we have $|a(s_{n_0+1})| \neq 1$ and as the case (8) is already discussed in Lemma 4.2 it follows with (6a) that it remains to consider

$$|a(s_{n_0+1})| > 1. \tag{4.2.14}$$

With (6a), (4) and (13) it is concluded

$$|a(s_{n_0})| < 1,$$

i.e. with $s = s_{n_0}$, $t = s_{n_0+1}$, (6), (8), (11) and (13) the assumptions of Corollary 4.1 in case 2 are fulfilled. It follows that (1) has no continuous solution on $[s_{n_0}, s_{n_0+1}]$ and consequently there exists no continuous solution of (1) defined on $\mathbf{R}^1$. Theorem 4.6 is thus demonstrated.

◊

In Theorem 2.1 and in case 1 of Theorem 4.6 the existence of a unique solution $f \in P$ is demonstrated. Contrary to Theorem 2.1, Assumptions 4.1 has to be satisfied in Theorem 4.6. However, the assumption

$$|a(x)| \neq 1, \forall x \in \mathbf{R}^1$$

is dropped. The necessary condition for ensuring a unique solution $f \in P$ is that the coefficient function a in (1) fulfils either $|a(s_n)| > 1, 0 \leq n \leq N - 1$

or $|a(s_n)| < 1, 0 \le n \le N-1$.

In the following the differentiability of f is investigated. Let P^m denote the set of the real-valued, m-times continuously differentiable 2π-periodic function defined on $\mathbf{R}^1$ and $f^{(k)}$, k = 1, 2,…., m with $f^{(1)} = f'$ the k^{th} derivative of a function $f \in P^m$. We differentiate (1) k-times and assuming $h'(x) > 0$ we find:

$$(f^{(k)} \circ h)(x) - A_k(x) \cdot f^{(k)}(x) = B_k(x),\ k = 1, 2, \ldots, m, \tag{4.2.15a}$$

where

$$A_k(x) = a(x) \cdot (h'(x))^{-k} \tag{4.2.15b}$$

and B_k is a function that depends on $f, \ldots, f^{(k-1)}, \ldots, f^{(1)} \circ h, \ldots, f^{(k-1)} \circ h$, $h^{(2)}, \ldots, h^{(k)}$, $a^{(1)}, \ldots, a^{(k)}$, $b^{(k)}$. However, B_k is independent of a and $f^{(k)}$. (15) is a linear functional equation of the form (1) for the unknown function $f^{(k)}$. Let $\mu_{s_n}^{(k)}$ be the characteristic exponent of (15) in the fixed point s_n. With (4) and (15) we have:

$$\mu_{s_n}^{(k)} = \frac{\log \dfrac{|a(s_n)|}{\left(h'(s_n)\right)^k}}{\log h'(s_n)} = \frac{\log|a(s_n)|}{\log h'(s_n)} - k = \mu_{s_n} - k \tag{4.2.16}$$

where $\mu_{s_n}^{(0)} = \mu_{s_n}$ and $k = 0, 1, 2, \ldots, m$, $0 \le n \le N-1$.

We consider the following notation. Let

$$\lceil c \rceil = \min_{\substack{z \in Z \\ z \ge c}} z$$

denote *the ceiling function* of a real number c.

The following theorem shows that from the assumptions $h(x) = x + p(x)$, $p \in P^m$, $a \in P^m$ and $b \in P^m$ it is not possible to conclude that f is an element of P^m. Relating to (17) in the following Theorem we note that the sign of μ_{s_n} is alternating. By Assumption 4.1 we have $N \ge 2$, i.e. at least one μ_{s_n} (n = 0,…, N) is positive. By using (15), (16), we find:

Theorem 4.7: Let the following assumptions hold for (1):
1. h and a satisfy Assumption 4.1;

2. $$\mu_{s_{n-1}} \cdot \mu_{s_n} < 0,\ 0 \le n \le N; \tag{4.2.17}$$

3. h, a, b $\in P^m$ where m is given by

$$m = \min\left\{ \mu_{s_n} \,\middle|\, \mu_{s_n} > 0, 0 \le n \le N \right\}. \tag{4.2.18}$$

With

$$M = \lceil m \rceil \tag{4.2.19}$$

we have for the solution f of (1):

1. $f \in P^{M-1}$;
2. If $B_M(x) \neq 0,\ \forall x \in \mathbf{R^1}$ then f is not in P^M.

Proof: Because of (17) and (19) we have M = 1, 2,…. For proving the first assertion we distinguish two cases:

1. M = 1
As the assumptions of Theorem 4.7 imply the assumption of Theorem 4.6, the assertion follows with Theorem 4.6.

2. M = 2, 3,....
The assertion is shown by induction on n. Let $f^{(k-1)} \in P$, $1 \le k \le M-1$ and it is to show $f^{(k)} \in P$. For k = 1 the assertion follows as in the case M = 1. Let $f^{(k-1)} \in P$, k = 2, 3,….. With assumption 2 we have

$$A_k \in P,\ B_k \in P \tag{4.2.20}$$

in (15). From (16), (17), (18) and (19) follows

$$\text{sign } \mu_{s_n}^{(k)} = \text{sign } \mu_{s_n},\ 0 \le k \le M-1,\ 0 \le n \le N.$$

Based on (17), (20) and assumption 2 of the theorem, it is concluded that, using $a(x) = A_k(x)$, $b(x) = B_k(x)$, $\forall x \in \mathbf{R^1}$, the functional equation (15) for $f^{(k)}$

fulfils the assumption of Theorem 4.6 in case 1, i.e. we find $f^{(k)} \in P$ and $f \in P^{M-1}$.

It remains to show assertion 2. From (16) and (18) it follows that there exists an index n_0, $0 \le n_0 \le N-1$ with

$$\mu_{s_{n_0}}^{(M)} \le 0,\ \mu_{s_{n_0+1}}^{(M)} < 0.$$

We consider 2 cases

1. $\mu_{s_{n_0}}^{(M)} < 0$

As the assumptions of case 2 of the theorem imply the assumption of case 1 it follows $f \in P$, $f \in P^1, \ldots, f \in P^{M-1}$. Using $\mu_{s_{n_0}+1}^{(M)} < 0$ and applying Theorem 4.6 in case 2 to (15), it follows that there exists no continuous solution on the interval $[s_{n_0}, s_{n_0+1}]$ thus $f \notin P^M$.

2. $\mu_{s_{n_0}}^{(M)} = 0$

Using the notation in (3.1.3c) we have

$$\mu_{s_{n_0}} = \frac{\log A_M(s_{n_0})}{\log h'(s_{n_0})} = \frac{\log \tilde{A}_M(0)}{\log \tilde{h}'(0)} = 0.$$

With Theorem 3.6 we conclude that the solution $\tilde{f}$ of (3.1.3b) has a logarithmic singularity in the origin, i.e. the solution is unbounded and with (3.1.3a) it follows that f^M is unbounded in s_{n_0}, i.e. $P \notin f^M$.

◊

From the second part of the proof of Theorem 4.7 it follows that $f^{(M)}$ is unbounded in s_n if $\mu_{s_n} = M$. This yields the following corollary:

Corollary 4.2: It is assumed that the coefficient functions of (1) satisfy the assumptions of Theorem 4.7 in case 2 and let $M \in \mathbf{N}$ be given by (18) and (19). If $\mu_{s_n} = M$ in s_n, $0 \le n \le N-1$ then

$$\left|f^{(M)}(x)\right| \to \infty, \quad |x| \to s_n,$$

i.e. $f^{(M)}$ is unbounded in s_n.

◊

The following corollary shows that even though the coefficient functions in (1) are assumed infinitely differentiable the solution f is not infinitely differentiable.

Corollary 4.3: Let the following assumptions hold for (1):
1. h and a satisfy Assumption 4.1;

2.
$$\mu_{s_{n-1}} \cdot \mu_{s_n} < 0, \ 0 \le n \le N;$$

3. h, a, b $\in P^{\infty}$.

By considering (18) and (19) there follows the assertions of Theorem 4.7 for the solution f of (1).

◊

Theorem 4.7, Corollary 4.2 and Corollary 4.3 show that in fixed points s_n with $\mu_{s_n} = m$ in general the solution of the functional equation (1) has a singularity in the M^{th}-derivative. In Section 4.4 the behaviour of these singularities in the Newton-Raphson process is further investigated.

4.3. The case without fixed points

As in Section 4.2 we consider

$$(f \circ h)(x) - a(x) \cdot f(x) = b(x) \tag{4.3.1}$$

for all $x \in \mathbf{R^1}$, where h: $\mathbf{R^1} \to \mathbf{R^1}$ with $h(x) = x + p(x)$, $p \in P$, $a \in P$ and $b \in P$ are given functions and f is unknown.

A. Rotation numbers

We start by focussing on h in (1). Contrary to the previous sections we do not assume that h has fixed points, i.e. that the iterates h^n, n = 1, 2, 3,…. of h are bounded. We use in this section some results from the literature on dynamical systems. The following lemma characterises the iterates of h and is a preparation for discussing further the solution f of (1).

Lemma 4.3: It is assumed that h: $\mathbf{R^1} \to \mathbf{R^1}$ is an invertible function $h(x) = x + p(x)$, $p \in P$. Then for an arbitrary $x_0 \in \mathbf{R^1}$ the sequence

$$x_n = \frac{h^n(x_0)}{n}, \; n = 1, 2, \ldots. \tag{4.3.2a}$$

converges. Furthermore the limit

$$\rho(h) = \frac{1}{2\pi} \lim_{n\to\infty} x_n \tag{4.3.2b}$$

is independent of the initial value x_0 and unique for h.

Proof: An intuitive and easy accessible demonstration of the Lemma is in Devaney [8, 2003], however, extensive discussions of the rotation number are also in i.e. Guckenheimer [6, 1983], De Melo [7, 1993], Herman [9, 1979].

◊

Definition 4.2: Following Lemma 4.3 ρ(h) is called the *rotation number of h.*

Remarks:

1. **(Fixed Points)** If h has fixed points the sequence $h^n(x_0)$, $n \to \infty$ in (2) is bounded and the evaluation of (2) yields $\rho(h) = 0$.

2. **(Translation)** Referring to Example 1.2 in Section 1.4 *the translation* s_τ, $\tau \in \boldsymbol{R}^1$ defined by

$$s_\tau(x) = x + \tau \tag{4.3.3}$$

is an example of a map that has for $\tau \neq 0$ no fixed points. For the rotation number we find with (2):

$$\rho(s_\tau) = \lim_{n \to \infty} \frac{x_0 + n\tau}{2\pi n} = \frac{\tau}{2\pi}. \tag{4.3.4}$$

As shown in the following the discussion of f in (1) requires distinguishing between irrational and rational rotation number.

3. **(Approximation of ρ)** In Chapter 8 we consider a one-parametric family of maps h and calculate the rotation number numerically.

4. **(Application to the solution of (1))** As shown in e.g. Guckenheimer [6, 1983] and Devaney [8, 2003] and applied in Paragraph D of this section the behaviour of the iterates of h is substantially different for rational and irrational rotation number. We see in this section that the same situation also applies to the discussion of the solution f of (1). The case ρ irrational is described in Theorem 4.10. We derive conditions under which the solution f of (1) is infinitely differentiable and Theorem 4.11 deals with $\rho \in \mathbf{Q}$.

B. The coefficient function a in (1)

As shown in Section 4.2 the discussion of (1) depends not only critically on h but also on a in (1). In Theorem 2.1 the convergence of the series (2.3.6) is concluded from $|a(x)| > 1, \ \forall x \in \mathbf{R}^1$. The following lemma is a preparation for showing the convergence of (2.3.6) under the more general condition (6).

We proceed by studying the product in series (2.3.6) and assume

$$\left|a(x)\right| > 0, \ \forall x \in \mathbf{R}^1 \tag{4.3.5}$$

in (1). We use this assumption, e.g. for applying the symmetry relation (2.3.22), (2.3.23) and (2.3.24) to (1) (see e.g. Theorem 4.9).

Lemma 4.4: Let the following assumptions hold for (1):
1. $a \in P$ with (5), $b \in P$;
2. $h(x) = x + p(x)$, $p \in P$.
If the convergence

$$\prod_{k=0}^{n} \left| \left(a \circ h^k\right)(x) \right|^{-1} \to 0, n \to \infty, \ \forall x \in \mathbf{R}^1 \tag{4.3.6}$$

is uniform, the series (2.3.6) is bounded by a real constant that is independent of $x \in \mathbf{R}^1$.

Proof: Since $[0, 2\pi]$ is compact, there follows from Assumption 1 that there exists a $C_a \in \mathbf{R}^1$ with $C_a > 0$ such that

$$\left|a(x)\right| > C_a > 0, \ \forall x \in [0, 2\pi],$$

and from 2π-periodicity we have

$$\left|a(x)\right| > C_a > 0, \ \forall x \in \mathbf{R}^1, \tag{4.3.7a}$$

hence the product

$$A_n(x) = \prod_{k=0}^{n-1} \left| \left(a \circ h^k\right)(x) \right|^{-1}, n = 0, 1, 2, \ldots, \ \forall x \in \mathbf{R}^1 \tag{4.3.7b}$$

is well-defined. Using the compactness of $[0, 2\pi]$, the assumption $b \in P$ and (2.3.12), the estimation of the right side of (2.3.6) yields:

$$\left| \sum_{n=0}^{\infty} \frac{\left(b \circ h^n\right)(x)}{\prod_{k=0}^{n} \left(a \circ h^k\right)(x)} \right| \le \sum_{n=0}^{\infty} \left| \left(b \circ h^n\right)(x) \right| A_{n+1}(x) \le \left\| b \right\| \sum_{n=0}^{\infty} A_{n+1}(x). \tag{4.3.8}$$

In order to simplify the proof, the term $A_0(x) = 1$ is added in the last sum of the right side (see (2.3.5)). It remains to show that the assertion holds for

$$\sum_{n=0}^{\infty} A_n(x).$$

From the uniform convergence of (6) that there exists a from x independent $N \in \mathbf{N}$ and a $\alpha \in \mathbf{R}^1$ with $0 < \alpha < 1$ such that

$$A_N(x) < \alpha. \tag{4.3.9}$$

From (7b) we have

$$A_{(j+1)N}(x) = (A_N \circ h^{jN})(x) \cdot A_{jN}(x),\ j = 0, 1, 2, \ldots .$$

As (9) is valid for arbitrary $x \in \mathbf{R}^1$, it follows:

$$A_{(j+1)N}(x) \le \alpha\, A_{jN}(x),\ j = 0, 1, 2, \ldots,$$

thus with $A_0(x) = 1$

$$A_{jN}(x) < \alpha^j,\ j = 0, 1, 2, \ldots . \tag{4.3.10}$$

Again using (7b) we find:

$$A_{jN+m}(x) = (A_{jN} \circ h^m)(x) \cdot A_m(x),\ m = 0, 1, \ldots, N-1,$$

i.e. with (10):

$$\sum_{j=0}^{\infty} A_{jN+m}(x) < \frac{A_m(x)}{1-\alpha},\ m = 0, 1, \ldots, N-1. \tag{4.3.11}$$

As $a \in P$ and (7a), it follows that the sum

$$C_{max} = C_0 + C_1 + C_2 + \ldots + C_{N-1}$$

Where

$$C_m = \max_{x \in \mathbf{R}^1} A_m(x),\ m = 0, 1, \ldots, N-1$$

is independent of $x \in \mathbf{R^1}$. The summation over m in (11) yields:

$$\sum_{m=0}^{N-1}\sum_{j=0}^{\infty} A_{jN+m}(x) < \frac{C_{max}}{1-\alpha}. \tag{4.3.12}$$

From (8) and (6) we have $A_n(x) > 0$, n = 0, 1, 2,…., i.e. the convergence in (12) is absolute and the sums can be exchanged:

$$\sum_{j=0}^{\infty}\sum_{m=0}^{N-1} A_{jN+m}(x) = \sum_{n=0}^{\infty} A_n(x) \le \frac{C_{max}}{1-\alpha}. \tag{4.3.13}$$

The lemma is thus demonstrated.

◊

C. The translation (3)

The imaginary unit is denoted with i. For $M \in \mathbf{R^1}$ with $M > 0$ we define a *strip* S_M with width M in **C** that contains the real axis by

$$S_M = \{z = x + i\,y \in \mathbf{C} \mid x \in \mathbf{R^1}, |y| \le M\}. \tag{4.3.14}$$

In the following we simplify the notation by suppressing the index M in S_M. Furthermore we define *the translation* s_τ on a strip S in **C** for $\tau \in \mathbf{R^1}$ by

$$s_\tau(z) = z + \tau,\ z \in S. \tag{4.3.15}$$

For $\tau \neq 0$ the translation (14) has no fixed points and the iterates

$$s_\tau^k(z) = z + k\,\tau,\ k = 0, 1, 2,\ldots. \tag{4.3.16a}$$

do not go outside S, i.e.

$$(z + k\,\tau) \in S,\ k = 0, 1, 2,\ldots.. \tag{4.3.16b}$$

In addition by evaluating (2) we find (4). We note that (15) is obviously the analytic extension of (3) to **C** and define as follows:

Definition 4.3: Let P^{ω} denote the set of real-analytic functions defined on $\mathbf{R^1}$ that for any f

1. f can be extended to an analytic function defined on the interior $\overset{o}{S}_M$ for some M = M(f) > 0.
2. f is 2π-periodic in the real part x of $z = x + i\, y \in S$.

We proceed by assuming that h in (1) is a translation (3) and following (16) we have control over the iterates in a strip S. Using Lemma 4.4 we prove:

Theorem 4.8: Let the following assumptions hold for (1):
1. $a \in P^{\omega}$ with (5), $b \in P^{\omega}$;
2. $h(x) = s_{\tau}(x) = x + \tau,\ \tau \neq 0$.

If the convergence

$$\prod_{k=0}^{n} |a(x+k\tau)|^{-1} \to 0, n \to \infty,\ \forall x \in \mathbf{R^1} \tag{4.3.17}$$

is uniform then equation (1) has a unique solution $f \in P^{\omega}$.

Proof: We show that series (2.3.6) is converging locally uniform and represents a function in P^{ω} that solves (1).

We follow the proof of Lemma 4.4 and start by considering (7b) for the translation (3):

$$A_n(x) = \prod_{k=0}^{n-1} |a(x+k\tau)|^{-1}. \tag{4.3.18}$$

As in (9) there follows from (17) that there exists a $N \in \mathbf{N}$ and a $\alpha \in \mathbf{R^1}$ with $0 < \alpha < 1$ such that

$$A_N(x) < \alpha. \tag{4.3.19}$$

For any $x \in [0, 2\pi)$ we introduce the sequence

$$x_k = s_{\tau}^{k}(x) \bmod 2\pi = (x + k\,\tau) \bmod 2\pi,\ k = 0, 1, 2, \ldots.$$

and the product

$$\widetilde{A}_N(x_k) = \prod_{k=0}^{N-1} |a(x_k)|^{-1}.$$

From the 2π-periodicity of a and consequently of the product (18) we conclude from (19)

$$\widetilde{A}_N(x_k) < \alpha. \tag{4.3.20}$$

We denote with R = [0, 2π] x [−M, M] a rectangle R with width M > 0. From the analyticity of a and b in a strip S of **C** and the compactness of R there follows by a continuity argument that there exists a rectangle with width M such that
1. (20) holds for all z in R, i.e.

$$\widetilde{A}_N(z) < \alpha,\ z \in R. \tag{4.3.21a}$$

2. b is bounded by a constant $C_b \in \mathbf{R}^1$ with $C_b > 0$

$$C_b = \sup_{z \in R} |b(z)|. \tag{4.3.21b}$$

Hence following the assumption of the 2π-periodicity in the definition of P^ω, we conclude that (21) is valid for a strip S in **C** with width M, i.e.

$$A_N(z) < \alpha,\ z \in S \tag{4.3.22a}$$

and

$$C_b = \sup_{z \in S} |b(z)|. \tag{4.3.22b}$$

In order to show locally uniform convergence of the series (2.3.6) we consider a z in S and a compact set K in an open neighbourhood of z. As in the proof of Lemma 4.4 there follows from (16) and (22) that the estimations (8) and (13) hold in S, i.e. the convergence of the series (2.3.6) is locally uniform and consequently (2.3.6) represents a real-analytic function f. The 2π-periodicity of f follows as in (2.3.11), i.e. we have $f \in P^\omega$. The uniqueness of f is concluded from (2.3.14), (16) and (22a).

◊

As preparation for Lemma 4.6 we need the following lemma:

Lemma 4.5: Let I = [u, v), v > u be a subinterval of [0, 1). We consider the translation (3) with rotation number (4). If $\rho(s_\tau)$ is irrational then the iterates starting with any x in [0, 1) of the translation s_τ are equally distributed modulo 2π, more precisely the ratio n_I of the iterates that lie in the interval I to the overall number n of iterates converges to the length v – u of the interval:

$$\lim_{n\to\infty} \frac{n_I}{n} = v - u .$$

Proof: see Weyl [31, 1916].

◊

Lemma 4.6: Let the following assumptions hold for (1):
1. $a \in P^\omega$ with (5);
2. $h(x) = s_\tau(x) = x + \tau$ with $\rho = \frac{\tau}{2\pi}$ irrational.
From

$$\int_0^{2\pi} \log |a(x)| dx > 0 \tag{4.3.23}$$

it is concluded that (17) holds.

Proof: From Assumption 1 of the lemma there follows that

$$\int_0^{2\pi} \log |a(x)| dx$$

can be approximated by Riemann sums from below. This is applied to (23), i.e. there exists a $N_0 \in \mathbf{N}$ and a partition

$$x_0 = 0 < \ldots\ldots < x_{N_0} = 2\pi$$

such that for refined partitions

$$x_0 = 0 < \ldots\ldots < x_N = 2\pi \tag{4.3.24}$$

with $N \geq N_0$ and

$$\{x_0, \ldots, x_{N_0}\} \subset \{x_0, \ldots, x_N\}$$

we have:

$$\int_0^{2\pi} \log|a(x)|\,dx > \sum_{j=0}^{N-1} m_j(x_{j+1} - x_j) > \frac{1}{2}\int_0^{2\pi} \log|a(x)|\,dx > 0, \qquad (4.3.25)$$

where

$$m_j = \inf\{\log|a(x)| \mid x \in [x_j, x_{j+1}]\},\ j = 0, 1, \ldots, N-1. \qquad (4.3.26)$$

For the rest of the proof we fix $N \geq N_0$ with (25). Let the sequence q_k be defined by

$$q_n(x) = s_\tau^n(x) \bmod 2\pi = (x + n\tau) \bmod 2\pi,\ n = 0, 1, 2, \ldots,\ \forall x \in [0, 2\pi)$$

and let n_j, $j = 0, \ldots, N-1$ denote the number of elements $q_n(x)$ that are in $[x_j, x_{j+1})$, $0 \leq j \leq N-1$. The total number of iterates n is then given by

$$n = \sum_{j=0}^{N-1} n_j\,.$$

We apply Lemma 4.5 to $[0, 2\pi)$:

$$2\pi \lim_{n\to\infty} \frac{n_j}{n} = d_j,\ j = 0, 1, \ldots, N-1,$$

where

$$d_j = x_{j+1} - x_j > 0,\ j = 0, 1, \ldots, N-1, \qquad (4.3.27)$$

i.e. for every $\delta > 0$ there exists a N_1 in $\mathbf{N}$ such that

$$\left| 2\pi\frac{n_j}{n} - d_j \right| < \delta,\ n \geq N_1,\ j = 0, 1, \ldots, N-1. \qquad (4.3.28)$$

(24) and (27) yield

$$d_j > 0,\ j = 0, 1, \ldots, N-1. \qquad (4.3.29)$$

As $\delta > 0$ in (28) is arbitrary we can choose $\delta \in \mathbf{R}^1$ such that

$$0 < \delta < \min_{0\leq j\leq N-1} d_j\,. \qquad (4.3.30)$$

With (28) we obtain

$$2\pi\frac{n_j}{n} > d_j - \delta > 0,\ n \geq N_1,\ j = 0, 1,\dots,N-1 \tag{4.3.31}$$

and, using (27) and (30), there exists $c_j > 0$ with

$$0 < c_j < \frac{d_j - \delta}{d_j},\ j = 0, 1,\dots, N-1. \tag{4.3.32a}$$

Furthermore we define $c \in \mathbf{R}^1$ with $c > 0$ by

$$c = 2\pi \min_{0\leq j\leq N-1} c_j. \tag{4.3.32b}$$

In the following equation (33) the sum of the left side is rearranged such that iterates

$$s_\tau^k(x) = x + k\,\tau,\ \forall\, x \in [0, 2\pi)$$

that are modulo 2π in the same interval $[x_j, x_{j+1})$, $j = 0, 1,\dots,N-1$ are also consecutive members in the sum. We consider the inverse of (17) and with (26), (31), (32) we have the following estimations:

$$\begin{aligned}
&\sum_{k=0}^{n} \log\left|a(x+k\tau)\right| \geq \sum_{j=0}^{N-1} m_j n_j = \\
&n\sum_{j=0}^{N-1} m_j \frac{n_j}{n} > \frac{n}{2\pi}\sum_{j=0}^{N-1} m_j(d_j - \delta) > \\
&\frac{n}{2\pi}\sum_{j=0}^{N-1} m_j c_j d_j > nc\sum_{j=0}^{N-1} m_j d_j,\ n > N_1.
\end{aligned} \tag{4.3.33}$$

Hence (25), (27) and (32) give

$$\sum_{k=0}^{n} \log\left|a(x+k\tau)\right| \to \infty,\ n \to \infty.$$

Taking the exponential yields:

$$\exp\left(\sum_{k=0}^{n} \log\left|a(x+k\tau)\right|\right) \to \infty,\ n \to \infty,$$

and considering the 2π-periodicity of a, we conclude (17) for any x in $\mathbf{R}^1$.

◊

Theorem 4.9: Let the following assumptions hold for (1):
1. $a \in P^{\omega}$ with (5), $b \in P^{\omega}$;
2. $h(x) = s_\tau(x) = x + \tau$ with $\rho = \frac{\tau}{2\pi}$ irrational.
If

$$\int_0^{2\pi} \log|a(x)|dx \neq 0, \tag{4.3.34}$$

then (1) has a unique solution $f \in P^{\omega}$.

Proof: Based on (34) we have to consider the following two cases:

1. $\int_0^{2\pi} \log|a(x)|\,dx > 0$

As the assumptions of the Theorem imply the assumption of Lemma 4.6 there follows (17) and consequently the assumptions of Theorem 4.8 are satisfied, i.e. we conclude that (1) has a unique solution $f \in P^{\omega}$.

2. $\int_0^{2\pi} \log|a(x)|\,dx < 0$

As $|a(x)| > 0$ by hypothesis, we consider (2.3.22) where

$$\tilde{h}(x) = h^{-1}(x) = s_\tau^{-1}(x) = x - \tau$$

and

$$\tilde{a}(x) = \frac{1}{a(x-\tau)},\ \tilde{b}(x) = -\frac{b(x-\tau)}{a(x-\tau)}.$$

Hence we have:

$$\int_0^{2\pi} \log|\tilde{a}(x)|dx = \int_0^{2\pi} \log\frac{1}{|a(x-\tau)|}dx = \int_0^{2\pi} \log\frac{1}{|a(x)|}dx > 0.$$

With case 1 it follows that (2.3.22) has a unique solution $f \in P^{\omega}$ and as the solutions of (2.3.22) are invariant under the transition from (2.3.22) to (1), equation (1) has also a unique solution $f \in P^{\omega}$.

◊

Considering Theorem 4.8 for coefficient functions in (1) that are in P, the following corollary holds:

Corollary 4.4: Let the following assumptions hold for (1):
1. $a \in P$ with (5) and $b \in P$;
2. $h(x) = s_\tau(x) = x + 2\pi\rho$ with ρ irrational.

If

$$\int_0^{2\pi} \log|a(x)|dx \neq 0$$

then (1) has a unique solution $f \in P$.

Proof: As in the proof of Theorem 4.9 we consider two cases:

1. $\int_0^{2\pi} \log|a(x)|\, dx > 0$

As the assumptions of the Theorem imply the assumptions of Lemma 4.6 it follows (17) and (6). Following Lemma 4.4 we conclude that the series (2.3.6) is converging uniformly and represents a continuous function f in P. The uniqueness of f follows as in the proof of Theorem 2.1. Case 1 of the corollary is thus demonstrated.

2. $\int_0^{2\pi} \log|a(x)|\, dx < 0$

The assertion follows as in the proof of case 2 in Theorem 4.9 by considering (2.3.22)-(2.3.24).

◊

Example 4.1: Let

$$a(x) = 1.5 - \cos x, \tag{4.3.35a}$$

$$b(x) = 1, \tag{4.3.35b}$$

$$h(x) = s_\tau(x) = x + \tau \tag{4.3.35c}$$

in (1). $\tau = 0$ yields

$$f(x) = \frac{1}{\cos x - 0.5}, x \neq \frac{\pi}{3} + 2k\pi, \ k \in \mathbf{Z} \tag{4.3.36}$$

and $\rho = 0$, thus

$$\rho \in \mathbf{Q}.$$

As the denominator in (36) vanishes for

$$x = \frac{\pi}{3} + 2k\pi, \ k \in \mathbf{Z}$$

we have

$$f \notin P.$$

However, for

$$\tau = \frac{1}{n}, n = 1, 2, 3, \ldots. \tag{4.3.37}$$

in (35c) we have

$$\rho = \frac{1}{2\pi n} \notin \mathbf{Q},$$

i.e. Assumption 2 of Theorem 4.9 is satisfied. Hence equation (1) with the coefficient functions (35) and (37) satisfies the assumptions of Theorem 4.9, i.e. there exists a unique solution $f \in P^\omega$.

◊

The discussion of a translation (3) for h in (1) is thus concluded and we are applying Corollary 4.4 in the following.

D. The general case

We further discuss the regularity of f in (1) and assume that $h(x) = x + p(x)$, $x \in \mathbf{R}^1$ in (1) is not necessarily a translation but an invertible function $h: \mathbf{R}^1 \to \mathbf{R}^1$. However, as the translation (3) is a special case of an invertible function h the following is concluded from Example 4.1:

1. We can not expect the regularity to be constant for rotation numbers that are in an open set of the real axis.
2. In order to discuss the regularity of the real-valued 2π-periodic solution f of (1) the cases rational and irrational rotation number have to be considered separately.

Referring to the shift of a fixed point of h in the origin in (3.1.3) our approach in the following is to transform h to a translation.

Definition 4.4: Let $\hat{h}$, $h: \mathbf{R}^1 \to \mathbf{R}^1$ be invertible maps with $h(x) = x + p(x)$, $p \in P$ and $\hat{h}(x) = x + \hat{p}(x)$, $\hat{p} \in P$. Then h and $\hat{h}$ are *conjugate* if and only if there exists an invertible map $D: \mathbf{R}^1 \to \mathbf{R}^1$ with $D(x) = x + d(x)$, $d \in P$ such that

$$\hat{h} = D \circ h \circ D^{-1}. \tag{4.3.38}$$

The following Lemma is a further result from the literature of dynamical systems. Under appropriate assumptions for h it ensures conjugacy (38) between h and the translation (3).

Lemma 4.7 (Denjoy): It is assumed that the derivative of the invertible function $h(x) = x + p(x)$, $p \in P^1$ has bounded variation and the rotation number $\rho(h)$ of h is irrational. Then the following holds:
1. There exists an invertible function $D: \mathbf{R}^1 \to \mathbf{R}^1$

$$D(x) = x + d(x),\ d \in P, \tag{4.3.39}$$

with

$$\hat{h}(x) = s_{2\pi\rho(h)}(x) = x + 2\pi\rho(h) \tag{4.3.40}$$

in (38), i.e. h is conjugate to the translation (3) with $\tau = 2\pi\rho(h)$. D is only determined up to an additive constant.

2. If we normalise with the mean

$$\int_0^{2\pi} d(x)\,dx = 0 \tag{4.3.41}$$

D(x) = x + d(x), d $\in$ P with (38) is unique. As in (2.3.18), there exists the representation

$$D^{-1}(x) = x + d_1(x) \tag{4.3.42}$$

where $d_1 = -\, d \circ D^{-1} \in$ P.

Proof: see i.e. De Melo [7, 1993], Herman [9, 1979].

◊

Remark: The regularity of D in (38) is extensive to discuss and to assess. We refer to Herman [9, 1979]. In Lemma 4.7 we only have continuity and not a precise degree of differentiability of D in (38).

Definition 4.5: We denote with P^{∞} the set of the real-valued, infinitely differentiable, 2π-periodic functions of a variable x $\in$ $\mathbf{R}^1$.

The following theorem is the main result of this section relating to irrational rotation number. It ensures a solution f of (1) in P^{∞}.

Theorem 4.10: Let the following assumptions hold for (1):

1. a $\in$ P^{∞} with (5), b $\in$ P^{∞};
2. h(x) = x + p(x), p $\in$ P^{∞} , x $\in$ $\mathbf{R}^1$ is invertible with h'(x) > 0;
3. There exists a D: $\mathbf{R}^1 \to \mathbf{R}^1$ such that D(x) = x + d(x) is invertible and satisfies (38), (39), (40) with an irrational rotation number ρ(h).

If

$$\frac{\int_0^{2\pi} \log\left|(a \circ D^{-1}(x))\right| dx}{\int_0^{2\pi} \log\,(h' \circ D^{-1})(x)\,dx} \neq 0,\ 1,\ 2, \ldots. \tag{4.3.43}$$

then equation (1) has a unique solution f $\in$ P^{∞}.

Proof: Let $f^{(k)}$ be the k^{th} derivative of f. We show by induction on k that (4.2.15) has a unique solution $f^{(k)} \in P$ for $k = 0, 1, 2, \ldots$.

Using Assumption 2 of the theorem it follows that Lemma 4.7 holds, i.e. it is concluded that there exists an invertible function D: $\mathbf{R^1} \to \mathbf{R^1}$ with

$$x = D(y),\ y \in \mathbf{R^1} \tag{4.3.44}$$

such that (39) and (40) hold. As $h'(x) \in P^{\infty}$ and $h'(x) > 0$ by hypothesis there exists a $C \in \mathbf{R^1}$ with $C > 0$ and

$$h'(x) > C > 0.$$

Consequently there follows $(h' \circ D^{-1})(y) > C > 0$, thus with (4.2.15) we have

$$f^{(k)} \circ h \circ D^{-1} - (A_k \circ D^{-1}) \cdot (f^{(k)} \circ D^{-1}) = B_k \circ D^{-1},\ k = 0, 1, 2, \ldots, \tag{4.3.45}$$

where

$$A_k = a \cdot (h')^{-k}.$$

Furthermore $B_k \in P$ in (45) is a function that is independent of a and $f^{(k)}$. Using the notation

$$f_1 = f^{(k)} \circ D^{-1} \tag{4.3.46}$$

in (45) implies:

$$f_1 \circ D \circ h \circ D^{-1} - (A_k \circ D^{-1}) \cdot f_1 = B_k \circ D^{-1},\ k = 0, 1, 2, \ldots. \tag{4.3.47}$$

Then, with Assumption 2 we find

$$f_1 \circ s_{2\pi\rho(h)} - (A_k \circ D^{-1}) \cdot f_1 = B_k \circ D^{-1},\ k = 0, 1, 2, \ldots. \tag{4.3.48}$$

where $s_{2\pi\rho(h)}$ is given by (40). Using (42) in Assumption 2, the assumption $f^{(m-1)} \in P$, $0 \le m \le k-1$ and the Assumption 1 of the theorem, it follows that
1. $A_k \circ D^{-1} \in P$ with $\left|(A_k \circ D^{-1})(x)\right| \neq 0,\ \forall x \in \mathbf{R^1}$,

2. $B_k \circ D^{-1} \in P$.
Furthermore we obtain

$$\int_0^{2\pi} \log\left|(A_k \circ D^{-1})(x)\right| dx = \int_0^{2\pi} \log \frac{\left|(a \circ D^{-1})(x)\right|}{((h' \circ D^{-1})(x))^k} dx =$$

$$\int_0^{2\pi} \log\left|(a \circ D^{-1})(x)\right| dx - k \int_0^{2\pi} \log (h' \circ D^{-1})(x) dx \; .$$

Assumption (43) of the theorem yields:

$$\int_0^{2\pi} \log\left|(A_k \circ D^{-1})(x)\right| dx \neq 0 .$$

Applying Corollary 4.4 to (47) with

$$a(x) = (A_k \circ D^{-1})(x),\ b(x) = (B_k \circ D^{-1})(x),\ \forall x \in \mathbf{R}^1$$

it is concluded with (48) that (47) has a unique solution $f_1 \in P$. From (39) and (42) in Lemma 4.7, (44) and (46) it follows that

$$f^{(k)} = f_1 \circ D$$

is also a unique solution $f^{(k)} \in P$ of (45) and, using (44), it is concluded that $f^{(k)} \in P$ in (4.2.15), i.e. we have

$$f \in P^{\infty}$$

for the solution f of (1).

◊

Remarks:

1. It is seen that the ratio in condition (43) is similar to the characteristic exponent defined in (4.2.4) and used in Section 4.2 for discussing the regularity of the solution f of (1). Contrary to Section 4.2, however, we have in Theorem 4.10 an integral condition for ensuring infinitely differentiability of f.

2. In Theorem 4.10 we only show infinitely differentiability and not analyticity of the solution f as we have no control over the iterates of h close to the real axis for h not being a translation (3). The question whether f can be continued to an analytic function into the complex plane is left to future research.

In the following it is assumed that the rotation number $\rho(h)$ is rational:

$$\rho(h) = \frac{p}{q}, \; p \in \mathbf{N_0}, \; q \in \mathbf{N_0}, \; p, q \text{ non divisible.}$$

For the iterates of s_τ we have

$$s_\tau^q(x) = x + 2\pi p\tau .$$

For distinguishing between irrational and rational rotation number we need:

Definition 4.6: We consider an invertible map h: $\mathbf{R^1} \to \mathbf{R^1}$ with $h(x) = x + p(x)$, $p \in P$. If there exists a $n = 0, 1, 2,\ldots.$ and a $x_0 \in \mathbf{R^1}$ such that $h^n(x_0) = x_0 \bmod 2\pi$ then x_0 is called a *periodic point* of h. As n is only unique up to a multiple in $\mathbf{Z}$, $N \in \mathbf{N_0}$ denotes the smallest number with $h^N(x_0) = x_0 \bmod 2\pi$ and N is called the *prime period* of x_0.

In addition $\mathbf{\Omega}$ denotes the set of periodic points $x \in \mathbf{R^1}$ of h and the sequences $h^n(x)$, $n = 0, 1, 2,\ldots,$ $\forall\; x \in \mathbf{\Omega}$ are called *periodic orbits.*

Remarks:

1. As the rotation number $\rho(h)$ is unique for a given h (compare Lemma 4.3) there follows for rational rotation number that the iterates of h are either periodic or tend to a periodic orbit and the prime period is unique.

2. In the case of a translation (3) we only have periodic orbits for rational rotation number.

The following Lemma characterises irrational rotation numbers (Devaney [8, 2003]):

Lemma 4.8: $\rho(h)$ is irrational if and only if h has no periodic points.

◊

The following theorem can be seen as a generalisation of Theorem 4.6 because fixed points are a special case of a periodic points ($h^0(x) = x$). Instead of conditions for the coefficient function a in the fixed points of h as discussed in Section 4.2, we have in Theorem 4.11 conditions for a and h that have to be valid for all periodic orbits of h.

Theorem 4.11: Let the following assumptions hold for (1):
1. $a \in P$ with (5), $b \in P$;
2. $h(x) = x + p(x)$ with $p \in P$ with $\rho(h) \in \mathbf{Q}$ and prime period N for all $x \in \mathbf{\Omega}$.
We distinguish two cases:

1. If

$$\prod_{k=0}^{N-1} \left| \left(a \circ h^k \right)(x) \right| > 1, \ \forall x \in \mathbf{\Omega} \tag{4.3.49}$$

the solution f of (1) is defined by series (2.3.6).

2. If

$$\prod_{k=0}^{N-1} \left| \left(a \circ h^k \right)(x) \right| < 1, \ \forall x \in \mathbf{\Omega} \text{ with h invertible,} \tag{4.3.50}$$

the solution f of (1) is defined by series (2.3.7) and in both cases equation (1) has a unique solution $f \in P$.

Proof: Let us first consider case 1, i.e. it is assumed that (49) holds. As the rotation number (2) is independent of x_0, it follows from the assumption $\rho \in \mathbf{Q}$ that the sequence $h^n(x_0)$, n = 0, 1, 2,…., $x_0 \in \mathbf{R^1}$ is either an element of a periodic orbit or converges towards a periodic orbit with prime period N, i.e. (49) implies (6). Using Lemma 4.4 it is concluded that the series (2.3.6) converges uniformly for $x \in \mathbf{R^1}$ and together with (2.3.8) and (2.3.11) we find that (2.3.6) represents a solution $f \in P$ of (1). In addition, (2.3.14) and (6) yield the uniqueness of f.

In order to prove case 2 it is assumed that (50) holds. As h is invertible by hypothesis and the function a does not vanish by Assumption 1, we consider (2.3.22). The coefficient functions $\tilde{a}$ and $\tilde{h} = h^{-1}$ satisfy

$$\prod_{k=0}^{N-1}\left|\left(\tilde{a}\circ\tilde{h}^{k}\right)(x)\right| = \prod_{k=0}^{N-1}\frac{1}{\left|\left(a\circ h^{-k}\right)(x)\right|} = \prod_{k=0}^{N-1}\frac{1}{\left|\left(a\circ h^{k}\right)(x)\right|} > 1,\ x\in\Omega$$

and with (2.3.18) it follows that there exists a function $p_1 = p \circ h^{-1} \in P$ with

$$\tilde{h}(x) = h^{-1}(x) = x + p_1(x),$$

therefore we find $\tilde{a} \in P$, $\tilde{b} \in P$. Thus, the coefficient functions of (2.3.22) satisfy the assumptions of case 1, i.e. (2.3.22) and (1) have a unique solution $f \in P$. By replacing the coefficient functions of (2.3.22) in (2.3.6), the representation (2.3.7) is easily verified.

◊

We illustrate Theorem 4.11 by two examples.

Example 4.2: Let

$$a(x) = \frac{3}{2},$$

$$b(x) = \sin x, \qquad (4.3.51)$$

$$h(x) = x + \pi + \frac{9}{8}\sin x + \frac{7}{16}\sin 2x$$

in (1). From

$$h^{n}(0) = (n+1)\pi,\ n = 0, 1, 2, \ldots \qquad (4.3.52)$$

there follows from the uniqueness of $\rho(h)$ and with (2)

$$\rho(h) = \frac{1}{2}$$

thus from Theorem 2.1 and as a in (51) is independent of h also from Theorem 4.11 it is concluded that there exists a unique solution $f \in P$ of (1) and using (2.3.6) with (51) we find $f(0) = 0$ and $f(\pi) = 0$. Derivation and iteration of (1) yield in $x = 0$

$$f'(h^2(0)) = \frac{a(h(0))\,a(0)}{h'(h(0))\,h'(0)} f'(0) + \frac{a(h(0))\,b'(0)}{h'(h(0))h'(0)} + \frac{b'(h(0))}{h'(h(0))}.$$

With (51) and (52) we have

$$f'(0) = f'(0) - \frac{2}{3}, \text{ i.e. } f' \notin P.$$

f is thus only continuous and not differentiable.

◊

Example 4.2 shows that in general from the assumptions $h \in P^{\omega}$, $a \in P^{\omega}$, $b \in P^{\omega}$ and $\rho(h) \in \mathbf{Q}$ cannot be concluded more regularity than continuity for the solution $f \in P$ of (1).

The example below shows that in general the assumption $\rho(h)$ irrational is necessary for showing the existence of a function

$$D(x) = x + d(x),\ d \in P \tag{4.3.53}$$

with (40) and (42).

Example 4.3: Let

$$a(x) = \frac{3}{2} - \sin x, \tag{4.3.54a}$$

$$b(x) = 1, \tag{4.3.54b}$$

$$h(x) = x + \frac{1}{2}\sin x \tag{4.3.54c}$$

in (1). Evaluating (4.2.4) yields

$$\mu_0 = \frac{\log 1.5}{\log 1.5} = 1,$$

$$\mu_\pi = \frac{\log 1.5}{\log 0.5} = -0.58,$$

thus (4.2.5) holds and it follows that the coefficient functions (54) of (1) satisfy the assumptions of Theorem 4.6, i.e. there exists a unique solution

$$f \in P \tag{4.3.55}$$

of (1). As h in (54c) has a fixed point the rotation number vanishes. Applying Lemma 4.7, it is concluded that h is conjugate to the identity, i.e we have $\hat{h}(x) = s_0(x) = x$ in (40). If (42) is assumed we obtain with k = 0 in (47)

$$f_1(y) = \frac{(b \circ D^{-1})(y)}{(a \circ D^{-1})(y) - 1}$$

where $x = D^{-1}(y)$. From (54a) it can be seen that $a(\frac{\pi}{6}) = 1$ and it follows that there exists $y_0 \in \mathbf{R^1}$ with $(a \circ D^{-1})(y_0) = 1$. Thus, we have $f_1 \notin P$ and, using (46) and assumption (53), we find

$$f = f_1 \circ D \notin P.$$

This is contradictory to (55), i.e. there exists no function (53).

◊

4.4. A convergence theorem

A. Introduction

Starting point for this section is the condition of invariance (1.1.7)

$$G(S(\varphi), \varphi) - S(H(S(\varphi), \varphi)) = 0, \varphi \in \mathbf{R^1}. \tag{4.4.1}$$

In Stiefel/Kirchgraber [12, 1978] and Nicolaisen [20, 1998] it is shown that under appropriate assumptions for the maps H and G and their derivatives there exists a unique continuously invariant curve S that satisfy (1). Both proofs are based on showing that iterating the condition of invariance (1) and applying Banach's fixed point theorem to an appropriate Banach Space yields the existence and uniqueness of a continuously function $S \in P$ that satisfies (1). Differentiation properties of invariant curves are also discussed. As iterating (1) and considering the Newton-Raphson method developed in this thesis are substantially different, the proofs in Stiefel/Kirchgraber and Nicolaisen cannot be applied to the Newton-Raphson methods considered in this work. In this section we present a convergence theorem relating to the iteration in Algorithm 1 (see Section 2.2).

As the Newton-Raphson method involves the derivative of the Newton-Raphson approximations the convergence theorems in this section assume that (1) has a continuously differentiable invariant curve S. As a consequence the coefficient functions of (2.2.2a) introduced in (2.2.2b-d) exist:

$$h(S)(\varphi) = H(S(\varphi), \varphi), \tag{4.4.2a}$$

$$a(S)(\varphi) = G_r(S(\varphi), \varphi) - S'(H(S(\varphi), \varphi)) \cdot H_r(S(\varphi), \varphi), \tag{4.4.2b}$$

$$b(S)(\varphi) = - S(H(S(\varphi), \varphi)) + G(S(\varphi), \varphi). \tag{4.4.2c}$$

As S satisfies the condition of invariance (1) we have

$$b(S)(\varphi) = 0, \varphi \in \mathbf{R^1} \tag{4.4.2d}$$

in (2c). In this section we show that under appropriate assumptions for the functions H and G the Newton-Raphson method is converging locally to the invariant curve, i.e. there exists an appropriate neighbourhood of S such that for every initial condition $S_0 \in P^\omega$ in this neighbourhood the iteration in Algorithm 1 (Section 2.2) is converging towards the invariant curve S. In Kantorovich [10, 1981] there are convergence theorems for Newton-Raphson iterations that we consider in Algorithm 2.2. However, they assume that the linear approximation is Fréchet differentiable (Definition see Section 2.1) and that the Newton-Raphson approximations map a given Banach Space into itself. As the linear problem in the Algorithm developed in this work requires the derivative of the Newton-Raphson approximations we only have differentiability in P^1 and not in P (Gâteaux differentiability). Consequently the convergence theorems in [10] are not directly applicable and we need a more specific analysis.

We recall the original Newton-Raphson process considered in (2.1.5). By first choosing an initial approximation $S_0 \in P^1$ for the invariant curve S we solve the linear functional equation

$$(d_n \circ h_n)(\varphi) - a_n(\varphi) \cdot d_n(\varphi) = b_n(\varphi),\ \varphi \in \mathbf{R}^1,\ n = 0, 1, 2, \ldots \tag{4.4.3a}$$

where d_n is unknown and the coefficient functions of (3a) are given by

$$h_n(\varphi) = h(S_n)(\varphi) = H(S_n(\varphi), \varphi), \tag{4.4.3b}$$

$$a_n(\varphi) = a(S_n)(\varphi) = G_r(S_n(\varphi), \varphi) - S_n'(H(S_n(\varphi), \varphi)) \cdot H_r(S_n(\varphi), \varphi), \tag{4.4.3c}$$

$$b_n(\varphi) = b(S_n)(\varphi) = -S_n(H(S_n(\varphi), \varphi)) + G(S_n(\varphi), \varphi). \tag{4.4.3d}$$

Consecutive Newton-Raphson approximations are then computed by

$$S_{n+1}(\varphi) = S_n(\varphi) + d_n(\varphi). \tag{4.4.3e}$$

By assuming a Newton-Raphson approximation S_n in P^1, n = 0, 1, 2,…., we have shown in Theorem 2.1 that assuming $h_n(\varphi) = \varphi + p_n(\varphi)$, $p_n \in P$, $a_n \in P$ with $|a_n(\varphi)| \neq 1$, $\forall \varphi \in \mathbf{R}^1$ and $b_n \in P$ the linear functional equation (3a) has a unique solution d_n in P and we distinguish the following two cases:

1. If $|a_n(\varphi)| > 1$, $\forall \varphi \in \mathbf{R}^1$, the solution d_n is defined by

$$d_n = -\sum_{j=0}^{\infty} \frac{b_n \circ h_n^j}{\prod_{i=0}^{j} a_n \circ h_n^i}. \tag{4.4.4}$$

2. If $|a_n(\varphi)| < 1$, $\forall \varphi \in \mathbf{R}^1$ and h_n is invertible with $\tilde{h}_n = h_n^{-1}$ the solution d_n is defined by

$$d_n = \sum_{j=0}^{\infty} b_n \circ \tilde{h}_n^{j+1} \prod_{i=1}^{j} a_n \circ \tilde{h}_n^i. \tag{4.4.5}$$

Summarizing it is seen that the preservation of the regularity in the Newton-Raphson process does not follow from Theorem 2.1, more precisely a *differentiable* Newton-Raphson approximation S_n in (3e) implies in general only a *continuous* Newton-Raphson correction d_n in (4), (5) respectively. The following two theorems reflect the fixed case discussed in Section 4.2 but also investigated in the convergence proof of this Section. Referring to Theorem 2.1, (4) and (5), they show that S_n in (3e) does not need to be differentiable and as a consequence a_n in (3c) does not need to be continuous in a fixed point in order to ensure continuity of the solution d_n in (3a).

As in Section 4.1 let [s, t], s, t $\in \mathbf{R}^1$, s < t be a closed interval on the real axis and C[s, t] denotes the set of the continuous, real-valued functions defined on [s, t]. We consider the restriction of (3a) to [s, t]

$$d_n(h_n(\varphi)) - a_n(\varphi) \cdot d_n(\varphi) = b_n(\varphi) \tag{4.4.6}$$

for all $\varphi \in [s, t]$. We denote with s^+ the limit from the right and with t^- the limit from the left.

Theorem 4.12 (continuous solution): Let in the n^{th} Newton-Raphson step, n = 0, 1, 2,…. the following assumptions hold for (6) for all $\varphi \in [s, t]$:
1. $h_n, b_n \in C[s, t]$;
2. $h_n(s) = s$, $h_n(\varphi) > \varphi$, $\forall \varphi \in (s, t)$, $h_n(t) = t$, h_n invertible on [s, t];
3. $a_n \in C(s, t]$, $|a_n(\varphi)| > 0$, $\forall \varphi \in [s, t]$, $|a_n(t)| > 1$;
4. $|a_n(\varphi)| \to \infty$, $\varphi \to s^+$.
Then there exists a unique solution $d_n \in C[s, t]$ of (6).

Proof: As [s, t] is compact it follows from Assumption 3 that there exists a $C \in \mathbf{R}^1$ with $C > 0$ and $a_n(\varphi) > C$, $\forall \varphi \in [s, t]$. Using Assumption 2 with $\tilde{h}_n = h_n^{-1}$ we consider the coefficient functions of (2.3.23) in (2.3.24):

$$\tilde{a}_n = \frac{1}{a_n \circ \tilde{h}_n}, \; \tilde{b}_n = -\frac{b_n \circ \tilde{h}_n}{a_n \circ \tilde{h}_n}.$$

Hence by Assumption 4 and continuity of $\tilde{h}_n$ we find

$$|\tilde{a}_n(\varphi)| \to 0, \; \varphi \to s^+ \text{ and } |\tilde{a}_n(t)| < 1.$$

From Theorem 4.3 ($a(t) < 1$) in case 1 we conclude that there exists a unique continuous solution d_n of (2.3.23) on [s, t]. As the transition from (6) to (2.3.22) leaves the solution of (6) invariant the assertion of the theorem is thus shown.

◊

From Theorem 4.12 we conclude:

Corollary 4.5 (continuous solution): Let in the n^{th} Newton-Raphson step, n = 0, 1, 2,…. the following assumptions hold for (6) for all $\varphi \in [s, t]$:
1. $h_n, b_n \in C[s, t]$;
2. $h_n(s) = s$, $h_n(\varphi) < \varphi$, $\forall \; \varphi \in (s, t)$, $h_n(t) = t$, h_n invertible on [s, t];
3. $a_n \in C[s, t)$, $|a_n(\varphi)| > 0$, $\forall \varphi \in [s, t]$, $|a_n(s)| > 1$;
4. $|a_n(\varphi)| \to \infty$, $\varphi \to t^-$.
Then there exists a unique solution $d_n \in C[s, t]$ of (6).

Proof: The assertion of the theorem follows from the proof of Theorem 4.4 because the mirrored equations (4.1.23) satisfy the assumption of Theorem 4.12.

◊

The following Theorem ensures solutions that are infinitely differentiable. Referring to Section 4.2 we have seen that the regularity of d_n is determined by the positive characteristics exponent and in Theorem 4.13 we have an infinitely differentiable solution of (6). In the following $C^\infty[s, t]$ denotes the set of the infinitely continuously differentiable functions defined on [s, t].

Theorem 4.13 (infinitely differentiable solution): Let in the n^{th} Newton-Raphson step, n = 0, 1, 2,…. the following assumptions hold for (6) for all $\varphi \in [s, t]$:
1. $h_n, a_n, b_n \in C^{\infty}[s, t]$;
2. $h_n(s) = s$, $h_n(\varphi) < \varphi$, $\forall \varphi \in (s, t)$, $h_n(t) = t$, h_n is invertible;
3. $a_n(s) = 0$, $|a_n(t)| < 1$.
Then there exists a unique solution $d_n \in C^{\infty}[s, t]$ of (6).

Proof: From Theorems 4.3 it follows that there exists a unique continuous solution on the interval [s, t]. Furthermore as it is assumed that the coefficient function h_n, a_n, b_n of (6) are infinitely differentiable we consider following (4.2.15) the k^{th} derivative of (6)

$$d_n^{(k)}(h_n(\varphi)) - A_{n,k}(\varphi) \cdot d_n^{(k)}(\varphi) = B_{n,k}(\varphi),\ k = 1, 2, \ldots \tag{4.4.7a}$$

where

$$A_{n,k}(\varphi) = a_n(\varphi) \cdot \left(\frac{\partial h_n(\varphi)}{\partial \varphi}\right)^{-k},\ \varphi \in [s, t] \tag{4.4.7b}$$

and $B_{n,k}$ is a function that depends on $d_n, \ldots, d_n^{(k-1)}, \ldots, d_n^{(1)} \circ h_n, \ldots, d_n^{(k-1)} \circ h_n$, $h_n^{(1)}, \ldots, h_n^{(k)}$, $a_n^{(1)}, \ldots, a_n^{(k)}$, $b_n^{(k)}$. However, $B_{n,k}$ is independent of a_n and $d_n^{(k)}$. Note that $A_{n,k}(s) = 0$ and using Assumption 1 and 2 of the theorem it is seen that the fixed point in t is repelling, i.e. we have $h_n^{(1)}(t) \geq 1$ and as a consequence $|A_{n,k}(t)| < 1$ in (7b). Applying Theorem 4.3 in case 1 to (7) it follows that there exists a unique solution $d_n^{(k)} \in C[s, t]$ of (7a) and we conclude $d_n \in C^{\infty}[s, t]$.

◊

B. The convergence of the Newton-Raphson approximations

In Section 4.2 the linear problem (3a) of the Newton-Raphson method has been studied if the invertible function $h_n(\varphi) = H(S_n(\varphi), \varphi)$, n = 0, 1, 2,…. has an even number $M \geq 2$ of fixed points. It is shown in Corollary 4.3 that even if we assume that the coefficient functions of the linear problem (3a) are infinitely differentiable the Newton-Raphson corrections are not infinitely differentiable, i.e. we loose differentiability in the Newton-Raphson iteration. Furthermore we have seen in Section 4.2 that the singularities of the Newton-Raphson approximations are in the fixed points with positive characteristic exponents (see definition in (4.2.4)).

In this section we assume that the circle map induced by the invariant curve (2a) has the same properties as the circle maps induced by the Newton-Raphson approximations (3e) with (3b) introduced in Assumption 4.1 and discussed in Section 4.2. The following Assumption 4.2 is used for showing convergence of the Newton-Raphson method (3) in Theorem 4.15. It ensures that the theorems of Sections 4.1 and 4.2 are applicable in each Newton-Raphson step. Circle maps induced by the invariant curve that have no fixed points are left to further research. However, as the properties (8) are preserved by iterating a circle map with (8) our convergence proof presented in this section is also applicable to the case of invariant curves with an induced circle map (2a) that have a rational rotation number. In Sections 8.2 (phase locking) and 9.2 (Model of delayed regulation) rational rotation numbers of circle maps are subject to further investigation. The examination of irrational rotation number of the circle map induced by the invariant curve is left to future research.

In order to keep control over the fixed points of H in the Newton-Raphson process we suppose in Assumption 4.2 and throughout in this section that the fixed points are invariant in the Newton-Raphson process. As illustrated by the test examples in Sections 6.2 and 6.3 and by the Model of delayed regulation (Section 9.2), fixed point on the invariant curve can often be computed explicitly by elementary numerically methods, i.e. for their computation is no need for the Newton-Raphson method studied in the thesis.

Assumption 4.2: The circle map (2a) induced by the invariant curve S and the circle map (3b) induced by the initial approximation $S_0 \in P$ of $S \in P$ are differentiable with

$$\frac{\partial H(S_0(\varphi),\varphi)}{\partial \varphi} > 0 \text{ and } \frac{\partial H(S(\varphi),\varphi)}{\partial \varphi} > 0, \forall\, \varphi \in \mathbf{R}^1. \tag{4.4.8a}$$

This implies in particular that they are invertible. Furthermore they are assumed to have an even number M of fixed points

$$s_m \in [0, 2\pi),\ m = 0, 1, \ldots, M-1,\ M \geq 2 \tag{4.4.8b}$$

such that

$$S_0(s_m) = S(s_m) \tag{4.4.8c}$$

i.e.

$$H(S(s_m), s_m) = H(S_0(s_m), s_m) \tag{4.4.8d}$$

with

$$\left.\frac{\partial S_0(\varphi)}{\partial \varphi}\right|_{\varphi=s_m} \neq \left.\frac{\partial S(\varphi)}{\partial \varphi}\right|_{\varphi=s_m} \tag{4.4.8e}$$

such that

$$\left.\frac{\partial H(S(\varphi), \varphi)}{\partial \varphi}\right|_{\varphi=s_m} \neq 1 \tag{4.4.8f}$$

and

$$\left.\frac{\partial H(r,\varphi)}{\partial r}\right|_{(r,\varphi)=(S(s_m),s_m)} \neq 0\,. \tag{4.4.8g}$$

The conditions for S and S_0 in Assumption 4.2 are similar to the assumptions used and discussed in the analysis of a Newton-Raphson process applied to a real-valued function of a real variable. In this section they are necessary for showing convergence of the Newton-Raphson process (3). Convergence proofs under weaker assumption than formulated in Assumption 4.2 and examination of the implied convergence speed are not subject of this thesis.

As illustrated in Chapters 6 and 9 in this work the above Assumption 4.2 is reflected and illustrated in various realisations of Algorithm 1:

1. The invariant curve computed in test example (6.2.2) with initial approximation (6.2.3) satisfies (8) although the circle map induced by (6.2.3) is the identity and thus consists only of fixed points. In particular in (8e), the examination of an initial approximation needs on the right side the first derivative of (1) with respect to φ. Our algorithm computes the invariant curve efficiently. We find convergence in a few Newton-Raphson steps (see Figures 4 and 5).

2. The model of delayed regulation (Section 9.2) shows that in the case of phase locking (Definition in Section 8.2) the fixed points of $H(S(\varphi), \varphi)$ on the invariant curve can be computed explicitly.
3. Following example (6.3.1) we have also convergence of Algorithm 1 in Section 2.2 if the fixed points (8b-d) on S and S_0 are close and we conjecture that our convergence proof can be extended to the case where the fixed point H do not coincide but are sufficiently close.

Lemma 4.6: From Assumption 4.2 there follows for n = 0, 1, 2,…. and m = 0, 1, 2,…., M – 1

$$H(S(s_m), s_m) = H(S_n(s_m), s_m), \tag{4.4.9a}$$

$$a(S_n)(s_m) = G_r(S_n(s_m), s_m) - S_n'(s_m) \cdot H_r(S_n(s_m), s_m), \tag{4.4.9b}$$

$$b(S_n)(s_m) = - S_n(H(S_n(s_m), s_m)) + G(S_n(s_m), s_m) = 0, \tag{4.4.9c}$$

$$S_{n+1}(s_m) = S_n(s_m). \tag{4.4.9d}$$

Proof: In the following we use induction with respect to the Newton-Raphson step n: For n = 0 the assertion (9a) is the same as (8b-d) and (9b) is concluded from (3c) and (9a). As S satisfy (1) we have from (1), (2c), (2d) and (8d)

$$b(S_0)(s_m) = - S_0(H(S_0(s_m), s_m)) + G(S_0(s_m), s_m) =$$

$$- S_0(s_m) + G(S_0(s_m), s_m) = - S(s_m) + G(S(s_m), s_m) =$$

$$- \ S(H(S(s_m), s_m)) + G(S(s_m), s_m) = b(S)(s_m) = 0.$$

As s_m is a fixed point of S_0 we have from (4) and (9c) that $d_0(s_m) = 0$, i.e. (9d) for n = 0 follows from (3e).

It is assumed that (9) is valid for n. From (9a) and (9d) we find

$$H(S(s_m), s_m) = H(S_n(s_m), s_m) = H(S_{n+1}(s_m), s_m) \tag{4.4.9e}$$

and from (9a) and (9d) there follows

$$a(S_{n+1})(s_m) = G_r(S_{n+1}(s_m), s_m) - S'_{n+1}(H(S_{n+1}(s_m), s_m)) \cdot H_r(S_n(s_m), s_m) =$$

$$G_r(S_{n+1}(s_m), s_m) - S'_{n+1}(s_m) \cdot H_r(S_{n+1}(s_m), s_m).$$

In addition we find with (9a), (9d), (9e) and the induction assumption

$$b(S_{n+1})(s_m) = - S_{n+1}(H(S_{n+1}(s_m), s_m)) + G(S_{n+1}(s_m), s_m) =$$

$$- S_{n+1}(s_m) + G(S_{n+1}(s_m), s_m) = - S_n(s_m) + G(S_n(s_m), s_m) = \qquad (4.4.9f)$$

$$- S_n(H(S_n(s_m), s_m)) + G(S_n(s_m), s_m) = b(S_n)(s_m) = 0.$$

Finally it follows by (4) and (9f) that $d_{n+1}(s_m) = 0$, i.e. we have with (3e) that

$$S_{n+2}(s_m) = S_{n+1}(s_m).$$

As we have shown the assertion of the Lemma for n + 1, the proof is completed.

◊

Summarising we have shown in Lemma 4.6 that the fixed points of H are invariant in the Newton-Raphson process and that the function values $S_n(s_m)$ are constant. Furthermore (9b) shows that we evaluate the derivative with respect to φ and as a consequence also $a(S_n)(s_m)$ in the fixed arguments s_m and additionally it is seen that (3d) vanishes in s_m in each Newton-Raphson step.

In the following we use some of the results in Sections 3.2, 4.1 and 4.2 to each Newton-Raphson step. More specifically we show that

1. The Newton-Raphson approximations S_n are converging locally on an interval containing an attractive fixed point s_m to the invariant curve S.
2. The Newton-Raphson approximations S_n are extended to elements of P.
3. The Newton-Raphson approximations S_n are converging in the Banach Space P towards a solution S of (1) in P.

For $M \in \mathbf{R}^1$ with $M > 0$ and $s \in \mathbf{R}^1$, $t \in \mathbf{R}^1$ with $s < t$ we define *a rectangle* R_M by

$$R_M = \{z = \varphi + i\,\rho \in \mathbf{C} \mid \varphi \in (s, t), \left|\rho\right| \le M\}.$$

Definition 4.7: Let $C^{\omega}(s, t)$ denote the set of real-analytic functions defined on (s, t) that for any f there exists some $M = M(f) > 0$ such that f is extendible to an analytic function on the interior $\overset{o}{R}_M$.

The following theorem is a preparation for showing the convergence of the Newton-Raphson method (3). As we consider in Theorem 4.15 analytic solutions of (3a), we need the analytic version of Theorem 4.1. The continuation of a continuous function by the linear functional equation is described in Theorem 4.1 and the following theorem essentially shows that an arbitrary analytic function can be continued uniquely analytically by (6).

Theorem 4.14: Let the following assumptions hold for (6) for all $\varphi \in [s, t]$:

1. $h_n, a_n, b_n \in C^{\omega}(s, t)$, $n = 0, 1, 2, \ldots$;
2. $h_n(s) = s$, $h_n(\varphi) < \varphi$, $\forall\, \varphi \in (s, t)$, $h_n(t) = t$, h_n invertible on [s, t];
3. For $\varphi_0 \in (s, t)$ let $f_l \in C^{\omega}(h_n(\varphi_0), \varphi_0)$ be a function that satisfies (6) in $\varphi = \varphi_0$:

$$f_l\,(h_n(\varphi_0)) = a_n(\varphi_0) \cdot f_l\,(\varphi_0) + b_n(\varphi_0).$$

Then f_l can be uniquely extended to $[h_n^j(\varphi_0), \varphi_0]$, $j = 1, 2, 3, \ldots$ by (6), i.e. there exists a unique function $f_j \in C[h_n^j(\varphi_0), \varphi_0]$ with $f_j \in C^{\omega}(h_n^j(\varphi_0), \varphi_0)$ such that

1. f_j satisfies (6), $\forall\, \varphi \in [h_n^{j-1}(\varphi_0), \varphi_0]$, $j = 1, 2, 3, \ldots$;
2. $f_j(\varphi) = f_l(\varphi)$, $\forall\, \varphi \in [h_n(\varphi_0), \varphi_0]$.

Proof: We follow the proof of Theorem 4.1 where we have shown that the continuations defined in (4.1.3) are continuous. Using the assumption of this theorem we conclude that the continuations are analytic with exception in $\varphi_{j-1} = h_n^{j-1}(\varphi_0)$, $j = 2, \ldots$. As shown in Theorem in 4.1 the continuations in φ_{j-1} are continuous but we conclude from the general theory of analytic functions that a function cannot be just continuous in an isolated point φ_{j-1} in the open set (φ_j, φ_0), but is analytically and uniquely extendible in φ_{j-1}.

◊

In Definition 1.3 neighbourhoods for a set of points close to the invariant curve S are introduced. For the following convergence proof, however, we need neighbourhoods of S that are defined for $v \in P$ and measured in the sup–norm.

Definition 4.8: It is assumed that S satisfies Assumption 1.3. Referring to (2.1.1) for $\varepsilon \in \mathbf{R^1}$ with $\varepsilon > 0$ a *neighbourhood* U_ε^{sup} of S is defined by $U_\varepsilon^{sup} = \left\{ v \in P \,\middle|\, \| v - S \|_{sup} < \varepsilon \right\}$.

In this Section we use the notation U_ε for U_ε^{sup}.

The existence and uniqueness of continuously invariant curves is investigated in Kirchgraber and Stiefel [13, 1978] and Nicolaisen [22, 1998]. We assume in the following the existence of a continuously differentiable invariant curve and our focus is to show convergence of the iteration (3) first proposed by Nicolaisen [22, 1998].

The following convergence proof pursues the structure of the previous sections: We first investigate the convergence locally in the neighbourhood of a fixed point (compare Sections 3.1 and 3.2) of the circle map induced by the invariant curve (see (2a)), followed by the examination of Newton-Raphson approximations defined on $\mathbf{R^1}$ (compare Sections 4.1 and 4.2).

Theorem 4.15 (Convergence Theorem): Following (1.1.4) we consider real-valued maps G and H of the real variables r and φ with the periodicity condition (1.1.5) that are analytic in r and φ. We assume the following:

1. Relating to Assumption 1.2 and 1.3 there exists a unique simply closed continuously differentiable curve parametrized by $r = S(\varphi)$ satisfying the condition of invariance (1).

2. The circle map (2a) induced by the invariant curve S satisfies Assumption 4.2.

3. In (2b) we have

$$a(S)(\varphi) > 0, \ \forall \varphi \in \mathbf{R^1}. \tag{4.4.10}$$

4. Referring to the fixed points s_m, $m = 0, 1, 2, \ldots, M-1$ introduced in Assumption 4.2 (see (8b-d)) we have either

$$a(S)(s_m) > 1, m = 0, 1, 2, \ldots, M - 1 \tag{4.4.11a}$$

or

$$0 < a(S)(s_m) < 1, m = 0, 1, 2, \ldots, M - 1. \tag{4.4.11b}$$

Then the Newton-Raphson approximations (3e) are continuous and the Newton-Raphson method (3) is locally converging, more precisely there exists a neighbourhood U_ε with $\varepsilon \in \mathbf{R}^1$ and $\varepsilon > 0$ of S such that for every initial approximation $S_0 \in P^\varpi$ (see Definition 4.3) in U_ε satisfying Assumption 4.2 the Newton-Raphson iteration (3) is converging in P towards the unique invariant curve $S \in P$ in U_ε.

Proof: Following (11a) and (11b) we have to consider the following two cases:

1. We first assume

$$a(S)(s_m) > 1, m = 0, 1, 2, \ldots, M - 1. \tag{4.4.12}$$

The assertion of the theorem is shown in 5 steps and in Step 1 to 4 we investigate solutions of (3a) on an interval

$$[s_{m-1}, s_{m+1}], m = 1, 3, \ldots, M - 1 \tag{4.4.13a}$$

where we consider with (4.2.3) and (8b) the 2π-periodic continuation of the fixed point s_0 given by

$$s_M = s_0 + 2\pi. \tag{4.4.13b}$$

Using Assumption 4.2 and in particular (8f) we assume without loss of generality that the fixed points are numbered such that

$$s_m = H(S(s_m), s_m) \tag{4.4.14a}$$

satisfies

$$0 < \left.\frac{\partial H(S(\varphi), \varphi)}{\partial \varphi}\right|_{\varphi = s_m} < 1, \tag{4.4.14b}$$

i.e. the fixed point s_m is attractive.

The overall idea of the following convergence proof is to find a subset of a neighbourhood of the invariant curve such that a Newton-Raphson process introduced in Algorithm 1 in Section 2.2 and explicitly formulated in (3) starting in this neighbourhood maps this subset into itself and converges in this neighbourhood to the invariant curve S.

(Perturbation in S): We have assumed in the theorem that h(S) in (2a) is continuous in S and as S is continuously differentiable by hypothesis it follows that a(S) in (2b) is continuous in S. As a consequence there follows from continuity by perturbation that there exists an $\varepsilon_1 \in \mathbf{R}^1$ with $\varepsilon_1 > 0$ and a neighbourhood U_{ε_1} of S such that for any $v \in P$ with v real-analytic on (s_{m-1}, s_{m+1}) for $m = 1,\ldots, M-1$ in U_{ε_1} the following holds:

1. As by Assumption 4.2 the circle map $H(S(\varphi), \varphi)$ induced by the invariant curve S is invertible, the circle map $H(v(\varphi), \varphi)$ induced by v is invertible.

2. As we have only continuity in the repelling fixed points by referring to the derivative in (2b) and assumption (10) of the theorem we have

$$a(v)(\varphi) > 0, \ \forall \varphi \in (s_{m-1}, s_{m+1}). \tag{4.4.15}$$

We proceed by defining the subset M_{ε_1} of the neighbourhood U_{ε_1} (Definition 4.8). The subset M_{ε_1} reflects the properties of the Newton-Raphson approximations to be considered in the following convergence proof: Analyticity in the attractive fixed points and continuity in the repelling fixed points. More specifically referring to Theorem 4.14 and to Step 4 of this convergence proof, the properties (16d), (16f) and (16g) in particular reflect that the Newton-Raphson approximations are in general only continuously extendible in the repelling fixed points. A function v belongs to the set M_{ε_1} if and only if the following holds:

1. $v \in U_{\varepsilon_1}$.

2. Referring to the fixed point s_m introduced in Assumption 4.2 and assumed in the convergence theorem the circle map H induced by v satisfies

$$v(s_m) = S(s_m), \tag{4.4.16a}$$

i.e.

$$s_m = H(v(s_m), s_m). \tag{4.4.16b}$$

3. We require

$$\frac{\partial H(v(\varphi), \varphi)}{\partial \varphi} > 0, \ \forall \varphi \ \in (s_{m-1}, s_{m+1}). \tag{4.4.16c}$$

and in the repelling fixed points v is continuous (m = 0, 2, 4,…., M). For the derivative of the circle function induced by v the limits

$$\lim_{\varphi \to s_m^+} \frac{\partial H(v(\varphi), \varphi)}{\partial \varphi} > 0, \ m = 0, 2, \ldots, M-2$$

and (4.4.16d)

$$\lim_{\varphi \to s_m^-} \frac{\partial H(v(\varphi), \varphi)}{\partial \varphi} > 0, \ m = 2, 4, \ldots., M.$$

exist and are finite.

4. In attractive fixed points we require

$$0 < \left. \frac{\partial H(v(\varphi), \varphi)}{\partial \varphi} \right|_{\varphi = s_m} < 1. \tag{4.4.16e}$$

5. By (16c) the one sided limits

$$a^-(v(s_m)) = \lim_{\varphi \to s_m^-} a(v(\varphi)), \ m = 0, 1, \ldots, M-1 \tag{4.4.16f}$$

and

$$a^+(v(s_m)) = \lim_{\varphi \to s_m^+} a(v(\varphi)), \ m = 1, 2, \ldots, M \tag{4.4.16g}$$

exist and it is assumed that v satisfies

$$a^-(v(s_m)) > 1, \ m = 0, 1, \ldots, M-1 \tag{4.4.16h}$$

and

$$a^{+}(v(s_m)) > 1, \; m = 1, 2, \ldots, M. \tag{4.4.16i}$$

We note that as S is continuously differentiable such functions v exist by (2b), (11a), (16a) and (16b).

As an example for (16a-e) we refer to Example 4.3. We proceed by choosing an arbitrary

$$v \text{ in } M_{\varepsilon_1}. \tag{4.4.17}$$

Moreover we consider neighbourhoods U_ε, $\varepsilon \in \mathbf{R^1}$ with $0 < \varepsilon < \varepsilon_1$ and corresponding subsets M_ε with $M_\varepsilon = U_\varepsilon \cap M_{\varepsilon_1}$ as specified by (16) for ε instead of ε_1.

Step 1 (the Newton-Raphson process in the fixed point s_m): In the following $D_\delta(s_m)$ denotes for $\delta \in \mathbf{R^1}$ with $\delta > 0$ the circle in **C** with radius δ centred in s_m. As the function a in (12), H in (14b) and H_r in (8g) are continuously differentiable in S and s_m by hypothesis there exists constants $C_1 \in \mathbf{R^1}$, $C_2 \in \mathbf{R^1}$, $C_3 \in \mathbf{R^1}$ with $C_1 > 0$, $C_2 > 0$, $C_3 > 0$, $\delta_1 \in \mathbf{R^1}$ with $\delta_1 > 0$ and a $\varepsilon_2 \in \mathbf{R^1}$ with $0 < \varepsilon_2 < \varepsilon_1$ such that for all $z \in \mathbf{C}$ in $D_{\delta_1}(s_m)$ and for all v in M_{ε_2} restricted to $D_{\delta_1}(s_m)$ we have

$$0 < C_1 < \left.\frac{\partial H(v(z),z)}{\partial z}\right|_{z=s_m} < C_2 < 1, \tag{4.4.18a}$$

$$a(v)(z) > C_3 > 1, \tag{4.4.18b}$$

and

$$\left.\frac{\partial H(r,\varphi)}{\partial r}\right|_{(r,\varphi)=(v(s_m),s_m)} \neq 0. \tag{4.4.18c}$$

As S is continuously real-differentiable by assumption of the theorem and $v \in P^{\omega}$ is assumed to be real-analytic in (s_{m-1}, s_{m+1}) the difference

$$\Delta_v(\varphi) = S(\varphi) - v(\varphi) \tag{4.4.18d}$$

of the restriction of S and v to $[s_m - \delta_1, s_m + \delta_1]$ is continuously real-differentiable.

In the following we derive by (28) a relationship between the first derivative $\Delta_v^{(1)}(\varphi)$ of the right side (18d) and the first derivative $b_v^{(1)}(\varphi)$ of

$$b_v(\varphi) = b(v)(\varphi) = -\, v(H(v(\varphi), \varphi)) + G(v(\varphi), \varphi) \tag{4.4.19}$$

in the fixed point s_m. Using (2d), (2.1.4) and we have

$$b(S)(\varphi) - b(v)(\varphi) + b(v)(\varphi) = 0.$$

As the residual $b_v(\varphi)$ on the invariant curve S vanishes in particular in the fixed points the evaluation $\varphi = s_m$ yields by (16a), (16b) and (8d)

$$b_v(s_m) = b(v)(s_m) = -\, v(H(v(s_m), s_m)) + G(v(s_m), s_m) =$$

$$-\, v(s_m) + G(v(s_m), s_m) = -\, S(s_m) + G(S(s_m), s_m) =$$

$$-\, S(H(S(s_m), s_m)) + G(S(s_m), s_m) = 0,$$

i.e.

$$b(v)(s_m) = 0. \tag{4.4.20}$$

We proceed by considering the linear functional equation

$$(d_v \circ h_v)(\varphi) - a_v(\varphi) \cdot d_v(\varphi) = b_v(\varphi),\ \varphi \in (s_m - \delta_1, s_m + \delta_1) \tag{4.4.21a}$$

where d_v is unknown and the coefficient functions of (21a) are given by (19) and

$$h_v(\varphi) = h(v)(\varphi) = H(v(\varphi), \varphi), \tag{4.4.21b}$$

$$a_v(\varphi) = a(v)(\varphi) = G_r(v(\varphi), \varphi) - v^{(1)}(H(v(\varphi), \varphi))\, H_r(v(\varphi), \varphi). \tag{4.4.21c}$$

As s_m is an attractive fixed point we conclude by (18b) and Theorem 2.1 that the series (4) converges and by using (20) we find

$$d_v(s_m) = 0. \tag{4.4.22}$$

Using (17) we consider the first derivative of (21a):

$$d_v^{(1)}(h_v(\varphi))\, h_v^{(1)}(\varphi) - a_v(\varphi)\, d_v^{(1)}(\varphi) - a_v^{(1)}(\varphi)\, d_v(\varphi) = b_v^{(1)}(\varphi). \tag{4.4.23a}$$

With (18a) it follows that the derivative of $h_v^{(1)}(\varphi)$ in s_m does not vanish and we have:

$$d_v^{(1)}(h_v(s_m)) - \frac{a_v(s_m)}{h_v^{(1)}(s_m)} d_v^{(1)}(s_m) - \frac{a_v^{(1)}(s_m)}{h_v^{(1)}(s_m)} d_v(s_m) = \frac{b_v^{(1)}(s_m)}{h_v^{(1)}(s_m)}, \tag{4.4.23b}$$

i.e. by using (18a), (18b) and (22)

$$d_v^{(1)}(s_m) = \frac{\dfrac{b_v^{(1)}(s_m)}{h_v^{(1)}(s_m)}}{1 - \dfrac{a_v(s_m)}{h_v^{(1)}(s_m)}}. \tag{4.4.24}$$

Referring to (27) the following considerations lead to a relationship between $b_v^{(1)}(s_m)$ and $\Delta_v^{(1)}(s_m)$. As S and the given maps G, H are assumed to be differentiable we derive the condition of invariance (1) with respect to φ

$$\begin{gathered} G_r(S(\varphi), \varphi)\, S^{(1)}(\varphi) + G_\varphi(S(\varphi), \varphi) - \\ S^{(1)}(H(S(\varphi), \varphi))\, [H_r(S(\varphi), \varphi)\, S^{(1)}(\varphi) + H_\varphi(S(\varphi), \varphi)] = 0. \end{gathered} \tag{4.4.25}$$

The first derivative of (18d)

$$S^{(1)}(\varphi) = v^{(1)}(\varphi) + \Delta_v^{(1)}(\varphi)$$

in (25) yields

$$G_r(S(\varphi), \varphi)\ (v^{(1)}(\varphi) + \Delta_v^{(1)}(\varphi)) + G_\varphi(S(\varphi), \varphi) -$$

$$(v^{(1)}(H(S(\varphi),\varphi)) + \Delta_v^{(1)}(H(S(\varphi),\varphi)))$$

$$[H_r(S(\varphi), \varphi)\ (v^{(1)}(\varphi) + \Delta_v^{(1)}(\varphi)) + H_\varphi(S(\varphi), \varphi)] = 0.$$

This is the same as

$$G_r(S(\varphi), \varphi)\ v^{(1)}(\varphi) + G_\varphi(S(\varphi), \varphi) -$$

$$v^{(1)}(H(S(\varphi),\varphi))\ [H_r(S(\varphi), \varphi)\ (v^{(1)}(\varphi) + \Delta_v^{(1)}(\varphi)) + H_\varphi(S(\varphi), \varphi)]$$

$$= - G_r(S(\varphi), \varphi)\ \Delta_v^{(1)}(\varphi) +$$

$$\Delta_v^{(1)}(H(S(\varphi),\varphi))\ [H_r(S(\varphi), \varphi)\ (v^{(1)}(\varphi) + \Delta_v^{(1)}(\varphi)) + H_\varphi(S(\varphi), \varphi)].$$

By rearranging such that the terms with $\Delta_v^{(1)}(\varphi)$ appear only on one side of the equation we have

$$G_r(S(\varphi), \varphi)\ v^{(1)}(\varphi) + G_\varphi(S(\varphi), \varphi) -$$

$$v^{(1)}(H(S(\varphi),\varphi))\ [H_r(S(\varphi), \varphi)\ v^{(1)}(\varphi) + H_\varphi(S(\varphi), \varphi)]$$

$$= - G_r(S(\varphi), \varphi)\ \Delta_v^{(1)}(\varphi) +$$

$$\Delta_v^{(1)}(H(S(\varphi),\varphi))\ [H_r(S(\varphi), \varphi)\ (v^{(1)}(\varphi) + \Delta_v^{(1)}(\varphi)) + H_\varphi(S(\varphi), \varphi)] +$$

$$v^{(1)}(H(S(\varphi),\varphi)\ H_r(S(\varphi), \varphi)\ \Delta_v^{(1)}(\varphi),$$

thus with the second order term at the end of the equation

$$G_r(S(\varphi), \varphi)\ v^{(1)}(\varphi) + G_\varphi(S(\varphi), \varphi) -$$

$$v^{(1)}(H(S(\varphi),\varphi))\ [H_r(S(\varphi), \varphi)\ v^{(1)}(\varphi) + H_\varphi(S(\varphi), \varphi)]$$

$$= - G_r(S(\varphi), \varphi)\ \Delta_v^{(1)}(\varphi) +$$

$$\Delta_v^{(1)}(H(S(\varphi),\varphi))\ [H_r(S(\varphi), \varphi)\ v^{(1)}(\varphi) + H_\varphi(S(\varphi), \varphi)] +$$

$$v^{(1)}(H(S(\varphi), \varphi))\ H_r(S(\varphi), \varphi)\ \Delta_v^{(1)}(\varphi) +$$

$$\Delta_v^{(1)}(H(S(\varphi), \varphi))\ H_r(S(\varphi), \varphi)\ \Delta_v^{(1)}(\varphi).$$

In the fixed point s_m we have with (16a)

$$G_r(v(s_m), s_m)\ v^{(1)}(s_m) + G_\varphi(v(s_m), s_m) -$$

$$v^{(1)}(H(v(s_m), s_m))\ [H_r(v(s_m), s_m)\ v^{(1)}(s_m) + H_\varphi(v(s_m), s_m)]$$

$$= - G_r(v(s_m), s_m)\ \Delta_v^{(1)}(s_m) +$$

$$\Delta_v^{(1)}(H(v(s_m), s_m))\ [H_r(v(s_m), s_m)\ v^{(1)}(s_m) + H_\varphi(v(s_m), s_m)] + \tag{4.4.26}$$

$$v^{(1)}(H(v(s_m), s_m))\ H_r(v(s_m), s_m)\ \Delta_v^{(1)}(s_m) +$$

$$\Delta_v^{(1)}(H(v(s_m), s_m))\ H_r(v(s_m), s_m)\ \Delta_v^{(1)}(s_m).$$

In the fixed point s_m we obtain using (16b)

$$\Delta_v^{(1)}(H(v(s_m), s_m)) = \Delta_v^{(1)}(s_m).$$

For the first two lines of equation (26) we refer to the first derivative of (18) and (25). Furthermore for the other side of equation (26) we refer to the first derivative of (21b) and (21c). As a consequence we find

$$b_v^{(1)}(s_m) = \left(-a_v(s_m) + h_v^{(1)}(s_m)\right)\Delta_v^{(1)}(s_m) +$$

$$\Delta_v^{(1)}(s_m)\ H_r(v(s_m), s_m)\ \Delta_v^{(1)}(s_m). \tag{4.4.27}$$

(24) yields

$$d_v^{(1)}(s_m) = \frac{\dfrac{\left(-a_v(s_m) + h_v^{(1)}(s_m)\right)}{h_v^{(1)}(s_m)}\Delta_v^{(1)}(s_m)}{1 - \dfrac{a_v(s_m)}{h_v^{(1)}(s_m)}} +$$

$$\frac{\dfrac{\Delta_v^{(1)}(s_m) H_r(v(s_m), s_m)\ \Delta_v^{(1)}(s_m)}{h_v^{(1)}(s_m)}}{1 - \dfrac{a_v(s_m)}{h_v^{(1)}(s_m)}}$$

$$=-\Delta_v^{(1)}(s_m)+\frac{\dfrac{\Delta_v^{(1)}(s_m)H_r(v(s_m),s_m)\,\Delta_v^{(1)}(s_m)}{h_v^{(1)}(s_m)}}{1-\dfrac{a_v(s_m)}{h_v^{(1)}(s_m)}}$$

$$=-\Delta_v^{(1)}(s_m)+\frac{\Delta_v^{(1)}(s_m)H_r(v(s_m),s_m)\,\Delta_v^{(1)}(s_m)}{h_v^{(1)}(s_m)-a_v(s_m)}. \qquad (4.4.28)$$

Summarizing by (22), (24), (27) and (28) we have derived relationships in the fixed point s_m and we proceed by investigating intervals that contain the fixed point s_m.

Step 2 (local convergence in s_m): From (18a), (18b) we conclude that for all z in $D_{\delta_1}(s_m)$ and for all v in M_{ε_2} restricted to $D_{\delta_1}(s_m)$ there exists a $M_1 \in \mathbf{R}^1$ defined by

$$M_1 = \min_{z\in D_{\delta_1}(s_m)} h_v^{(1)}(z)$$

such that

$$0 < M_1 < 1 \qquad (4.4.29a)$$

and a $M_2 \in \mathbf{R}^1$ defined by

$$M_2 = \min_{z\in D_{\delta_1}(s_m)} \frac{a_v(z)}{h_v^{(1)}(z)}$$

such that we have

$$M_2 > 1. \qquad (4.4.29b)$$

Furthermore as the derivative of the difference (18d) is continuous in a neighbourhood of (S, s_m) and as $\Delta_S^{(1)}(s_m) = 0$ by assumption of this Theorem it can be chosen arbitrarily small in the neighbourhood of (S, s_m) and as $H(v(\varphi), \varphi)$ satisfies (8g), (18c) respectively there exists an $\varepsilon_3 \in \mathbf{R}^1$ with $0 < \varepsilon_3 < \varepsilon_2$, a M_{ε_3} of S and a $\kappa \in \mathbf{R}^1$ with $0 < \kappa < 1$ such that for the restriction of v to $[s_m - \delta_1, s_m + \delta_1]$ we have

$$\left| H_r(v(s_m), s_m)\,\Delta_v^{(1)}(\varphi) \right| < \kappa \cdot (M_2 - 1)\, M_1,$$
$$\forall v \in M_{\varepsilon_3},\ \forall \varphi \in [s_m - \delta_1, s_m + \delta_1] \qquad (4.4.29c)$$

where using the notation in (18d) we have

$$\Delta_v^{(1)}(\varphi) = S^{(1)}(\varphi) - v^{(1)}(\varphi). \qquad (4.4.30)$$

Following (23a) we proceed by considering

$$\hat{b}_v^{(1)}(\varphi) = a_v^{(1)}(\varphi)\, d_v(\varphi) + b_v^{(1)}(\varphi) \qquad (4.4.31a)$$

and relating to (27) we make the ansatz

$$\hat{b}_v^{(1)}(\varphi) = \beta_v(\varphi)\Delta_v^{(1)}(\varphi) + \gamma_v(\varphi)\Delta_v^{(1)}(\varphi),\ \varphi \in [s_m - \delta_1, s_m + \delta_1]. \qquad (4.4.31b)$$

In this relationship $\hat{b}_v^{(1)}(\varphi)$ is given by (31a) and $\Delta_v^{(1)}(\varphi)$ is given by (30). We define $\gamma_v(\varphi)$ by

$$\gamma_v(\varphi) = \gamma(v)(\varphi) = H_r(v(s_m), s_m)\,\Delta_v^{(1)}(\varphi). \qquad (4.4.31c)$$

In (31b) the function $\beta_v(\varphi)$ is determined by solving for $\beta_v(\varphi)$. For $\varphi \to s_m$ we have in s_m by (27) for v = S

$$\beta_S(s_m) = h_S^{(1)}(s_m) - a_S(s_m). \qquad (4.4.31d)$$

As $\beta_S(s_m)$ is continuous in S and s_m there exists a $\delta_2 \in \mathbf{R}^1$ with $0 < \delta_2 < \delta_1$ such that $\beta_v(\varphi)$ is continuous for $\varphi \in [s_m - \delta_2, s_m + \delta_2]$.

Referring to (8e) in Assumption 4.2 we consider a v in M_{ε_3} that satisfy

$$v^{(1)}(s_m) \neq S^{(1)}(s_m). \qquad (4.4.32)$$

We proceed by considering the function $\alpha_v(\varphi)$ on $\varphi \in [s_m - \delta_2, s_m + \delta_2]$ defined by

$$\alpha_v(\varphi) = \alpha(v)(\varphi) = \sum_{j=0}^{\infty} \frac{\dfrac{\beta_v(h_v^j(\varphi))}{h_v^{(1)}(h_v^j(\varphi))}\,\Delta_v^{(1)}(h_v^j(\varphi))}{\prod\limits_{i=0}^{j} \dfrac{a_v(h_v^i(\varphi))}{h_v^{(1)}(h_v^i(\varphi))}}. \qquad (4.4.33a)$$

In the fixed point s_m we have by summing up the geometric series and using (31c) with (27)

$$\alpha_v(s_m) = -\Delta_v^{(1)}(s_m) .$$

Considering estimation (2.3.13) and (29b) there follows that the series in (33a) represents a continuous function in a neighbourhood of s_m and we conclude that there exists a $\varepsilon_4 \in \mathbf{R}^1$ with $\varepsilon_4 > 0$ and

$$\varepsilon_4 < \min \{\varepsilon_3, 1-\kappa \},$$

a $\delta_3 \in \mathbf{R}^1$ with $0 < \delta_3 < \delta_2$ and a function $r_v(\varphi) = r(v)(\varphi)$ with

$$\alpha_v(\varphi) = (-1 + r_v(\varphi))\ \Delta_v^{(1)}(\varphi),\ \forall \varphi \in [s_m - \delta_3, s_m + \delta_3],\ \forall v \in M_{\varepsilon_4} \tag{4.4.33b}$$

such that

$$\left| r_v(\varphi) \right| < \frac{\varepsilon_4}{2} \tag{4.4.33c}$$

and

$$r_v(s_m) = 0.$$

The investigations for an arbitrary v in the neighbourhood of s_m satisfying (32) are thus concluded.

In the following Assertion 1 in (34a) implies the local convergence in s_m and Assertion 2 in (34b) ensures condition (32) in the Newton-Raphson iteration more specially we claim: Considering the first derivative of (3e) with respect to φ and assuming an initial condition $S_0 \in P^\omega$ in M_{ε_4} the sequence $S_n^{(1)}(\varphi)$ is such that

1. there exists some $q \in \mathbf{R}^1$ with $0 < q < 1$ and some $\delta \in \mathbf{R}^1$ with $\delta > 0$ such that restrictions of $S_n^{(1)}(\varphi)$, n = 0, 1, 2,…. to $[s_m - \delta, s_m + \delta]$ satisfies

$$\left| S_n^{(1)}(\varphi) - S^{(1)}(\varphi) \right| < q^n \left| S_0^{(1)}(\varphi) - S^{(1)}(\varphi) \right|,\ \forall \varphi \in [s_m - \delta, s_m + \delta]. \tag{4.4.34a}$$

2. we have

$$S_n^{(1)}(s_m) - S^{(1)}(s_m) \neq 0 . \tag{4.4.34b}$$

We start by showing assertion (34) for n = 0 and consider the restriction of the first derivative of an arbitrary initial approximation $S_0 \in P^{\varpi}$ satisfying Assumption 4.2 in M_{ε_4} to $[s_m - \delta_3, s_m + \delta_3]$. Following (3e) we consider the first derivative of the iteration (3e) for n = 0

$$S_1^{(1)}(\varphi) - S^{(1)}(\varphi) = S_0^{(1)}(\varphi) - S^{(1)}(\varphi) + d_0^{(1)}(\varphi),\ \forall\, \varphi \in [s_m - \delta_3, s_m + \delta_3].$$

By using (30) with $v = S_0$ and $v = S_1$ this is the same as

$$\Delta_1^{(1)}(\varphi) = \Delta_0^{(1)}(\varphi) - d_0^{(1)}(\varphi).$$

As $\hat{b}_v^{(1)}(\varphi)$ in (31a) is analytic in s_m and as the constants (29a) and (29b) are valid on $D_{\delta_3}(s_m)$ we conclude that Theorem 3.4 is applicable to (23b) on $D_{\delta_3}(s_m)$, i.e.

$$d_0^{(1)}(z) = -\sum_{j=0}^{\infty} \frac{\dfrac{\hat{b}_0^{(1)}(h_0^j(z))}{h_0^{(1)}(h_0^j(z))}}{\prod\limits_{i=0}^{j} \dfrac{a_0(h_0^i(z))}{h_0^{(1)}(h_0^i(z))}},\ \forall\, z \in D_{\delta_3}(s_m) \qquad (4.4.35)$$

represents a real-analytic function. We proceed by considering the real variable φ

$$\Delta_1^{(1)}(\varphi) = \Delta_0^{(1)}(\varphi) + \sum_{j=0}^{\infty} \frac{\dfrac{\hat{b}_0^{(1)}(h_0^j(\varphi))}{h_0^{(1)}(h_0^j(\varphi))}}{\prod\limits_{i=0}^{j} \dfrac{a_0(h_0^i(\varphi))}{h_0^{(1)}(h_0^i(\varphi))}},\ \forall\, \varphi \in [s_m - \delta_3, s_m + \delta_3]$$

hence by (31b) and (33a) with $v(\varphi) = S_0(\varphi)$

$$\Delta_1^{(1)}(\varphi) = \Delta_0^{(1)}(\varphi) + \alpha(S_0)(\varphi) + \sum_{j=0}^{\infty} \frac{\dfrac{\gamma(S_0)(h_0^j(\varphi))}{h_0^{(1)}(h_0^j(\varphi))}\,\Delta_0^{(1)}(h_0^j(\varphi))}{\prod\limits_{i=0}^{j} \dfrac{a_0(h_0^i(\varphi))}{h_0^{(1)}(h_0^i(\varphi))}}$$

we find with (33b)

$$\Delta_1^{(1)}(\varphi) = r(S_0)(\varphi)\,\Delta_0^{(1)}(\varphi) + \sum_{j=0}^{\infty} \frac{\dfrac{\gamma(S_0)(h_0^j(\varphi))}{h_0^{(1)}(h_0^j(\varphi))}\,\Delta_0^{(1)}(h_0^j(\varphi))}{\prod\limits_{i=0}^{j} \dfrac{a_0(h_0^i(\varphi))}{h_0^{(1)}(h_0^i(\varphi))}},$$

by the triangle inequality and (29a) we have

$$\left|\Delta_1^{(1)}(\varphi)\right| \le \left|r(S_0)(\varphi)\right|\left|\Delta_0^{(1)}(\varphi)\right| + \left|\frac{1}{M_1}\sum_{j=0}^{\infty} \frac{\gamma(S_0)(h_0^j(\varphi))\,\Delta_0^{(1)}(h_0^j(\varphi))}{\prod\limits_{i=0}^{j} \dfrac{a_0(h_0^i(\varphi))}{h_0^{(1)}(h_0^i(\varphi))}}\right|$$

hence by using the definition of γ_v with $v(\varphi) = S_0(\varphi)$ in (31c)

$$\left|\Delta_1^{(1)}(\varphi)\right| \le \left|r(S_0)(\varphi)\right|\left|\Delta_0^{(1)}(\varphi)\right| +$$

$$\left|\frac{1}{M_1}\sum_{j=0}^{\infty} \frac{H_r(S_0(s_m),s_m)\,\Delta_0^{(1)}(h_0^j(\varphi))\Delta_0^{(1)}(h_0^j(\varphi))}{\prod\limits_{i=0}^{j} \dfrac{a_0(h_0^i(\varphi))}{h_0^{(1)}(h_0^i(\varphi))}}\right|.$$

The maximum on the right side yields

$$\left|\Delta_1^{(1)}(\varphi)\right| \le \sup_{\varphi\in[s_m-\delta_3,s_m+\delta_3]} \left|r(S_0)(\varphi)\right|\left|\Delta_0^{(1)}(\varphi)\right| +$$

$$\frac{1}{M_1}\sup_{\varphi\in[s_m-\delta_3,s_m+\delta_3]}\left|\sum_{j=0}^{\infty} \frac{H_r(S_0(s_m),s_m)\,\Delta_0^{(1)}(h_0^j(\varphi))\Delta_0^{(1)}(h_0^j(\varphi))}{\prod\limits_{i=0}^{j} \dfrac{a_0(h_0^i(\varphi))}{h_0^{(1)}(h_0^i(\varphi))}}\right|,$$

hence by the triangle inequality and estimation (29c)

$$\left|\Delta_1^{(1)}(\varphi)\right| \le \sup_{\varphi\in[s_m-\delta_3,s_m+\delta_3]} \left|r(S_0)(\varphi)\right|\left|\Delta_0^{(1)}(\varphi)\right| +$$

$$\kappa \sum_{j=0}^{\infty} \frac{(M_2 - 1)}{\prod_{i=0}^{j} \left| \dfrac{a_0(h_0^i(\varphi))}{h_0^{(1)}(h_0^i(\varphi))} \right|} \left| \Delta_0^{(1)}(\varphi) \right|.$$

This yields with (29b) and (33c)

$$\left| \Delta_1^{(1)}(\varphi) \right| < \left(\frac{\varepsilon_4}{2} + \kappa \right) \left| \Delta_0^{(1)}(\varphi) \right|$$

where ε_4 and κ are independent from n and with

$$0 < q = \frac{\varepsilon_4}{2} + \kappa < 1$$

we find

$$\left| \Delta_1^{(1)}(\varphi) \right| < q \left| \Delta_0^{(1)}(\varphi) \right|.$$

As a consequence we have

$$\left| S_1^{(1)}(\varphi) - S^{(1)}(\varphi) \right| < q \left| S_0^{(1)}(\varphi) - S^{(1)}(\varphi) \right|, \ \forall \varphi \in [s_m - \delta_3, s_m + \delta_3].$$

We conclude that the distance from $S_1^{(1)}(\varphi)$ to the first derivative of the invariant curve $S^{(1)}(\varphi)$ is diminishing on $[s_m - \delta_3, s_m + \delta_3]$, i.e. the restriction of

$$S_1^{(1)}(\varphi)\Big|_{[s_m - \delta_3, s_m + \delta_3]} \text{ is in } U_{\varepsilon_4}\Big|_{[s_m - \delta_3, s_m + \delta_3]}.$$

Hence from (9a) (Invariance of the fixed point (16a-b)), (18a) and (18b) (we refer to (16d-i)) we conclude

$$S_1^{(1)}(\varphi)\Big|_{[s_m - \delta_3, s_m + \delta_3]} \text{ is in } M_{\varepsilon_4}\Big|_{[s_m - \delta_3, s_m + \delta_3]},$$

therefore we have (34a) for n = 1. In addition we conclude (34b) for $\Delta_1^{(1)}(s_m)$ from the assumption $\Delta_0^{(1)}(s_m) \neq 0$, (28) and (18c).

As the estimations in (29) and (33) are valid for any restrictions v in M_{ε_4} to $[s_m - \delta_3, s_m + \delta_3]$ and as $S_n^{(1)}(\varphi)$ is assumed to be in M_{ε_4} with (32) and (34b) by induction assumption we note that it follows as for n = 1

$$\left| S_{n+1}^{(1)}(\varphi) - S^{(1)}(\varphi) \right| < q \left| S_n^{(1)}(\varphi) - S^{(1)}(\varphi) \right|, \ \forall \varphi \in [s_m - \delta_3, s_m + \delta_3]$$

with a constant q independent from n. Then we conclude by induction

$$\left|S_{n+1}^{(1)}(\varphi) - S^{(1)}(\varphi)\right| < q^{n+1} \left|S_0^{(1)}(\varphi) - S^{(1)}(\varphi)\right|, \forall \varphi \in [s_m - \delta_3, s_m + \delta_3].$$

For n + 1 the assertion 1 in (34a) with $\delta \in \mathbf{R}^1$ and $\delta = \delta_3 > 0$ is shown and in addition we conclude (34b) for $\Delta_{n+1}^{(1)}(\varphi)$ from the assumption of the induction, (18c) and (28). The proof of (34) is thus concluded.

As $\left|S_0^{(1)}(\varphi) - S^{(1)}(\varphi)\right|$ is considered on $[s_m - \delta, s_m + \delta]$ we conclude that the convergence is independent of φ and consequently the assertion (34) is satisfied for any $\delta \in \mathbf{R}^1$ with $\delta = \delta_3$ and in addition we have

$$\lim_{n\to\infty} S_n^{(1)}(\varphi) = S^{(1)}(\varphi), \ \forall \varphi \in [s_m - \delta, s_m + \delta].$$

The convergence is uniform on $[s_m - \delta, s_m + \delta]$ and we conclude from elementary analysis that

$$\lim_{n\to\infty} S_n(\varphi) = S(\varphi), \ \forall \varphi \in [s_m - \delta, s_m + \delta] \tag{4.4.36a}$$

is continuous. Furthermore considering (3e) we have

$$\lim_{n\to\infty} d_n(\varphi) = 0, \ \forall \varphi \in [s_m - \delta, s_m + \delta] \tag{4.4.36b}$$

hence from continuity and (2c), (2d), (3d) respectively we conclude

$$\lim_{n\to\infty} b_n(\varphi) = 0, \ \forall \varphi \in [s_m - \delta, s_m + \delta]. \tag{4.4.36c}$$

Step 3 (Newton-Raphson approximations are continuous on $[s_{m-1}, s_{m+1}]$ and analytic on (s_{m-1}, s_{m+1})): Following (13) we consider an interval $[s_{m-1}, s_{m+1}]$ and Newton-Raphson approximations S_n, n = 0, 1, 2,…. in M_{ε_4}. By induction we show that the restriction of S_n to $[s_{m-1}, s_{m+1}]$ satisfies

$$S_n \in C[s_{m-1}, s_{m+1}] \text{ with } S_n \in C^{\omega}(s_{m-1}, s_{m+1}), n = 0, 1, 2,\ldots.. \tag{4.4.37}$$

For n = 0 the assertion follows from the choice of the initial condition $S_0 \in P^{\omega}$ in the Theorem and we assume (37) for arbitrary n. Following Theorem 3.4 there exists in the fixed point s_m an analytic solution d_n in a neighbourhood of s_m. As $S_n \in M_{\varepsilon_4}$ it follows by (16d) and thus applying (2.3.22), (2.3.23) and assumption (16f), (16h) to Theorem 4.3 and Theorem 4.14 that there exists a unique solution

$$d_n(\varphi) \in C^{\omega}(s_{m-1}, s_m) \tag{4.4.38}$$

and continuously extendible in s_{m-1} and s_m.

On the interval $[s_m, s_{m+1}]$, we conclude by (16d) and thus considering (2.3.22), (2.3.23), (16g), (16i) from Theorem 4.4 that there exists a unique solution $d_n \in C[s_m, s_{m+1}]$. From the proof of Theorem 4.4 it follows from $d_n \in C^{\omega}(s_{m-1}, s_m)$ and (4.1.23) - (4.1.27) that $d_n \in C^{\omega}(s_m, s_{m+1})$. Finally in the fixed point s_m there exists locally a unique analytic solution of (6) (Theorem 3.4) and as the solution of (6) on $[s_{m-1}, s_m]$ and $[s_m, s_{m+1}]$ is also unique, we conclude (37) for n + 1 by considering (3e). By referring to (38) we proceed by investing a rectangle in $\mathbf{C}$ on which d_n is real-analytic for $n \to \infty$.

Step 4 (Convergence of Newton-Raphson method on $[s_{m-1}, s_{m+1}]$): Referring to Step 2, we consider a neighbourhood $D_\delta(s_m)$ with $\delta = \delta_3$ in the attractive fixed point s_m and as H in (2a) is invertible by (8a) the sequence

$$\varphi_j = h^{-j}(S)(\varphi_0) = H^{-j}(S(\varphi_0), \varphi_0),\ j = 0, 1, 2, \ldots \tag{4.4.39}$$

with $\varphi_0 \in \mathbf{R^1}$, $\varphi_1 \in \mathbf{R^1}$ such that

$$\varphi_0 \in (s_m - \delta, s_m],\ h(S)(\varphi_0) \in (s_m - \delta, s_m] \text{ and } \varphi_1 \in [s_{m-1}, s_m - \delta]$$

exists.

As the function a in (12) and H in (8f), (14b) are continuously differentiable in S and s_{m-1} by hypothesis there exists constants $C_4 \in \mathbf{R^1}$, $C_5 \in \mathbf{R^1}$ with $C_4 > 0$, $C_5 > 0$ and $\delta_4 \in \mathbf{R^1}$ with $\delta_4 > 0$ such that for all φ in $[s_{m-1}, s_{m-1} + \delta_4)$ and for all v in M_{ε_4} restricted to $[s_{m-1}, s_{m-1} + \delta_4)$ we have

$$\left.\frac{\partial H(v(\varphi),\varphi)}{\partial \varphi}\right|_{\varphi = s_{m-1}} > C_4 > 1$$

and

$$a(v)(\varphi) > C_5 > 1.$$

As s_{m-1} is repelling there exists a $N \in \mathbf{N}$ in (39) such that

$\varphi_N \in [s_{m-1}, s_{m-1} + \delta_4)$ and $\varphi_{N-1} \notin [s_{m-1}, s_{m-1} + \delta_4)$.

As discussed in Sections 3.1 and 4.1 the continuations process into fixed points of the circle maps (3b) of solutions d_n, n = 0, 1, 2,…. in (3a) (see e.g. Theorem 4.2) stops and moreover the solutions of (3a) are in general not analytic in such fixed points. Following (29a) it follows that s_m is a isolated fixed point in $D_\delta(s_m)$ and it is seen by (40) below that by Assumption 4.2 than there exists a rectangle in **C** containing $[\varphi_N, h(\varphi_0)]$ with no complex fixed points, i.e. the analytic Newton-Raphson corrections d_n, n = 0, 1, 2,…. over $[\varphi_N, h(\varphi_0)]$ can be continued analytically in a rectangle in **C** containing $[\varphi_N, h(\varphi_0)]$. We proceed by the circle map induced by the invariant curve (2a) and use the presentation (see (8a), (2.3.1), (2.3.18) and (1.1.5))

$$H(S(\varphi), \varphi) = \varphi + p(S(\varphi), \varphi), \ \forall \varphi \in \mathbf{R}^1$$

where p is continuously differentiable with

$$p(S(\varphi + 2\pi), \varphi + 2\pi) = p(S(\varphi), \varphi), \ \forall \varphi \in \mathbf{R}^1$$

by hypothesis of this Theorem. From Assumption 4.2 it follows that the restriction of p to $[\varphi_N, h(\varphi_0)]$ has no fixed points, hence we have

$$\left| p(S(\varphi), \varphi) \right| > 0, \ \forall \varphi \in [\varphi_N, h(\varphi_0)].$$

By continuity there exists a $\rho_1 \in \mathbf{R}^1$ with $\rho_1 > 0$, a rectangle

$$R_{\rho_1} = \{z = \varphi + i\,\rho \in \mathbf{C} \mid \varphi \in [\varphi_N, h(\varphi_0)], \left|\rho\right| \le \rho_1\}$$

and a $\varepsilon_5 \in \mathbf{R}^1$ with $0 < \varepsilon_5 < \varepsilon_4$ such that the restrictions of all $v \in M_{\varepsilon_5}$ to $[\varphi_N, h(S)(\varphi_0)]$ satisfy

$$\left| p(v(z), z) \right| > 0, \ \forall z \in R_{\rho_1}, \tag{4.4.40}$$

i.e. there are no fixed points in R_{ρ_1} and on $[\varphi_N, s_m]$ fixed points are not accumulating.

We continue by constructing a rectangle that overlaps $D_{\delta_4}(s_{m-1})$. As described in Step 2 of this proof and in particular in (35) the Newton-Raphson corrections $d_n(z)$, n = 0, 1, 2,…. are real-analytic in $D_\delta(s_m)$ and with

$$\rho_2 = \delta - s_m + \varphi_0$$

we have

$$d_n(z) = \sum_{j=0}^{\infty} d_{n,j}(\varphi_0)(z-\varphi_0)^j,\ d_{n,j} \in \mathbf{R^1},\ \forall\, z \in \overset{o}{D}_{\rho_2}(\varphi_0), \qquad (4.4.41a)$$

where $\overset{o}{D}_{\rho_2}(\varphi_0)$ denotes the interior of $D_{\rho_2}(\varphi_0)$. Thus it is concluded with (36b)

$$d_{n,j}(\varphi_0) \to 0,\ n \to \infty,\ j = 0, 1, 2, \ldots. \qquad (4.4.41b)$$

and let $\rho_3 \in \mathbf{R^1}$ with

$$\rho_3 = \min\{\rho_1, \rho_2\}.$$

By introducing the sequence $q_0 = \varphi_0, q_1, \ldots, q_K$ with $K \in \mathbf{N}$

$$q_j = q_0 - j \cdot \frac{\rho_3}{2},\ j = 1, \ldots, K \text{ and } q_{K+1} \le \varphi_N \text{ and } q_K > \varphi_N, \qquad (4.4.42)$$

the circles with centres q_j, $j = 1, \ldots, K$ with radius ρ_3 and the circle with centre $\hat{q}_k = \varphi_N + \rho_3$ and radius ρ_3 we consider the rectangle

$$R_{\rho_3} = \{z = \varphi + i\,\rho \in \mathbf{C} \mid \varphi \in [\varphi_N + \frac{\rho_3}{2}, h(\varphi_0)],\ |\rho| \le \frac{\sqrt{3}}{2}\rho_3\}.$$

Geometrically the rectangle R_{ρ_3} is contained in the circles considered with radius ρ_3 above, but $\varphi_N + \frac{\rho_3}{2}$ is not necessarily in $(s_{m-1}, s_{m-1} + \delta_4)$. We consider the circle with radius

$$\rho_4 = \frac{s_{m-1} + \delta_4 - \varphi_N}{2} \text{ in } D_{\delta_4}(s_{m-1})$$

and $\rho_5 \in \mathbf{R^1}$ with

$$\rho_5 = \min\{\rho_4, \rho_3\}.$$

Then with $\hat{\varphi}_N = \varphi_N + \frac{\rho_5}{2}$ the rectangle

$$R_{\rho_5} = \{z = \varphi + i\,\rho \in \mathbf{C} \mid \varphi \in [\hat{\varphi}_N, h(\varphi_0)],\ |\rho| \le \frac{\sqrt{3}}{2}\rho_5\}$$

is completely within $D_{\delta_4}(s_{m-1})$ on the left side.

We consider an initial condition S_0 in M_{ε_5} such that the restriction of S_0 to $[s_{m-1}, s_{m+1}]$ satisfies $S_0 \in C^{\omega}(s_{m-1}, s_{m+1})$ and $S_0 \in C[s_{m-1}, s_{m+1}]$ and we continue by illustrating the circle chain analytic continuation from $(\varphi_0, s_m]$ to $(\varphi_2, s_m]$. We note that the sequence (39) converges to s_{m-1}, i.e. we have $\varphi_2 \in (s_{m-1}, s_m)$. From (36) it follows

$$\lim_{n\to\infty} S_n(\varphi) = S(\varphi),\ \forall \varphi \in [\varphi_0, s_m], \tag{4.4.43a}$$

$$\lim_{n\to\infty} d_n(\varphi) = 0,\ \forall \varphi \in [\varphi_0, s_m] \tag{4.4.43b}$$

and

$$\lim_{n\to\infty} b_n(\varphi) = 0,\ \forall \varphi \in [\varphi_0, s_m]. \tag{4.4.43c}$$

By (41) we have

$$d_n(z) = \sum_{j=0}^{\infty} d_{n,j}(\varphi_0)(z-\varphi_0)^j,\ d_{n,j} \in \mathbf{R}^1,\ \forall z \in \overset{o}{D}_{\rho_5}(\varphi_0) \tag{4.4.44a}$$

and

$$d_{n,j}(\varphi_0) \to 0,\ n \to \infty,\ j = 0, 1, 2, \ldots.. \tag{4.4.44b}$$

By (42) it follows that there exists $k \in \mathbf{N}$ with $k \le K$ and $q_0 = \varphi_0, q_1, \ldots, q_k$ such that

$$q_{k+1} \le \varphi_2 \text{ and } q_k > \varphi_2.$$

In addition we consider the circle with centre $\hat{q}_k = \varphi_2 + \rho_5$ and radius ρ_5. From Theorem 4.14 we conclude that the Newton-Raphson corrections $d_n(z)$, $n = 0, 1, 2, \ldots.$ can be extended to $C^{\omega}(\varphi_2 + \frac{\rho_5}{2},\ s_m)$ defined on a rectangle R_M with a width of at least

$$M = M(d_n) = \frac{\sqrt{3}}{2}\rho_5, \; n = 0, 1, 2, \ldots..$$

We conclude by (43):

$$\lim_{n\to\infty} d_n(\varphi) = 0, \;\; \forall \varphi \in (\varphi_2 + \frac{\rho_5}{2}, s_m]. \qquad (4.4.45a)$$

We consider by (4.4.4):

$$d_n(\varphi) = -\sum_{j=0}^{\infty} \frac{b_n(h_n^j(\varphi))}{\prod_{i=0}^{j} a_n(h_n^i(\varphi))} =$$

$$-\sum_{j=1}^{\infty} \frac{b_n(h_n^j(\varphi))}{\prod_{i=0}^{j} a_n(h_n^i(\varphi))} - \frac{b_n(\varphi)}{a_n(\varphi)}, \;\; \forall \varphi \in (\varphi_1, \varphi_0].$$

Applying (45a) and (43c) we find

$$\lim_{n\to\infty} b_n(\varphi) = 0, \;\; \forall \varphi \in (\varphi_1, \varphi_0]$$

and

$$\lim_{n\to\infty} b_n(\varphi) = 0, \;\; \forall \varphi \in (\varphi_2 + \frac{\rho_5}{2}, \varphi_1],$$

i.e. we have (43c)

$$\lim_{n\to\infty} b_n(\varphi) = 0, \;\; \forall \varphi \in (\varphi_2 + \frac{\rho_5}{2}, s_m]. \qquad (4.4.45b)$$

Repeating the consideration from $(\varphi_0, s_m]$ to $(\varphi_2, s_m]$ we expand the convergence to $(\hat{\varphi}_N, s_m]$ and find by (45)

$$\lim_{n\to\infty} d_n(\varphi) = 0, \;\; \forall \varphi \in (\hat{\varphi}_N, s_m]$$

and

$$\lim_{n\to\infty} b_n(\varphi) = 0, \;\; \forall \varphi \in (\hat{\varphi}_N, s_m]. \qquad (4.4.46)$$

There exists a $\varepsilon_6 \in \mathbf{R}^1$ with $0 < \varepsilon_6 < \varepsilon_5$ and there exists a $n_0 \in \mathbf{N}$ such that

$$\left|d_n(\varphi)\right| < \varepsilon_6, \ \forall n > n_0, \ \forall \varphi \in [\hat{\varphi}_N, s_m],$$

$$\left|b_n(\varphi)\right| < \varepsilon_6, \ \forall n > n_0, \ \forall \varphi \in [\hat{\varphi}_N, s_m].$$

As $M_{\varepsilon_6} \subset M_{\varepsilon_5}$ we have

$$a(v)(s_{m-1}) > 1, \ \forall v \in M_{\varepsilon_6}.$$

By (16a), (16g), (16i) and applying (2.3.22), (2.3.23) to Theorems 4.2 and 4.3 yield

$$\lim_{\varphi \to s_{m-1}^+} d_n(\varphi) = d_n(s_{m-1}) = \frac{b_n(s_{m-1})}{1 - a_n^+(s_{m-1})}, \ \forall n > n_0. \tag{4.4.47}$$

With (9c) we have

$$b_n(s_{m-1}) = 0, \ \forall n > n_0 \tag{4.4.48}$$

and by (47) it follows

$$d_n(s_{m-1}) = 0, \ \forall n > n_0.$$

Hence by (46) and (48)

$$\lim_{n \to \infty} b_n(\varphi) = 0, \ \forall \varphi \in [s_{m-1}, s_m].$$

From the assumption in this Theorem that S is a unique solution of

$$G(S(\varphi), \varphi) - S(H(S(\varphi), \varphi)) = 0, \ \forall \varphi \in \mathbf{R}^1$$

it is implied that S is unique in

$$G(S(\varphi), \varphi) - S(H(S(\varphi), \varphi)) = 0, \ \forall \varphi \in [s_{m-1}, s_m]$$

and with (39) in

$$G(S(\varphi), \varphi) - S(H(S(\varphi), \varphi)) = 0, \ \forall \varphi \in (\varphi_j, s_m], j = 0, 1, 2, \ldots.. \quad (4.4.49)$$

By (43a) we find for all $(\varphi_1, \varphi_0]$

$$\lim_{n\to\infty} G(S_n(\varphi), \varphi) = \lim_{n\to\infty} S_n(H(S_n(\varphi), \varphi))$$

$$= \lim_{n\to\infty} S_n(H(S_n(\varphi), \varphi)) - \lim_{n\to\infty} S(H(S_n(\varphi), \varphi)) + \lim_{n\to\infty} S(H(S_n(\varphi), \varphi)).$$

From

$$\lim_{n\to\infty} S_n(H(S_n(\varphi), \varphi)) - \lim_{n\to\infty} S(H(S_n(\varphi), \varphi)) = 0$$

we have

$$\lim_{n\to\infty} G(S_n(\varphi), \varphi) = S(H(S(\varphi), \varphi)).$$

By continuity of G in the first argument we have

$$G(\lim_{n\to\infty} S_n(\varphi), \varphi) = S(H(S(\varphi), \varphi)), \ \forall \ \varphi \in (\varphi_1, \varphi_0].$$

As a consequence $\lim_{n\to\infty} S_n(\varphi)$ is the continuation of S to $(\varphi_1, \varphi_0]$ and thus by uniqueness of S in (49)

$$\lim_{n\to\infty} S_n(\varphi) = S(\varphi), \ \forall \ \varphi \in (\varphi_1, s_m].$$

Repeating the argument from $(\varphi_0, s_m]$ to $(\varphi_1, s_m]$ yields

$$S(\varphi) = \lim_{n\to\infty} S_n(\varphi), \ \forall \varphi \in [s_{m-1}, s_m]. \quad (4.4.50)$$

The proof for $(s_m, s_{m+1}]$ follows similarly the proof for $[s_{m-1}, s_m)$ due to the fact that the iterates of $h_n(\varphi)$ are separated by the two intervals and in a neighbourhood of s_m we have a unique real-analytic solution in each Newton-Raphson step (Theorem 3.4).

Summarizing we have shown that the Newton-Raphson iteration (3) is converging on each interval $[s_{m-1}, s_{m+1}]$, m = 1, 2,…., M – 1.

Step 5 (Convergence of Newton-Raphson method on $\mathbf{R}^1$): As we have only used Theorem 4.2 for continuation in the repelling fixed points the functions (3c) might not be continuous in the these fixed points. However, by Step 4 and more specifically by (46) and (47), we have with (16f)-(16i) in the fixed point s_m

$$d_n(s_m) = \frac{b_n(s_m)}{1 - a_n^+(s_m)} = \frac{b_n(s_m)}{1 - a_n^-(s_m)} = 0, \; m = 2, \ldots, M-2,$$

$$d_n(s_0) = \frac{b_n(s_0)}{1 - a_n^+(s_0)} = 0, \; d_n(s_M) = \frac{b_n(s_M)}{1 - a_n^-(s_M)} = 0,$$

i.e. the Newton-Raphson corrections d_n are continuous on $[s_0, s_M]$ and following (48) we have convergence of the Newton-Raphson process for all $\varphi \in [s_0, s_M]$. Referring to (4.2.3) in Section 4.2 and the 2π-periodicity of the coefficient functions in (3a) there exists a unique 2π-periodical continuous continuation $[s_0, s_M]$ of the Newton-Raphson corrections d_n and of the Newton-Raphson approximations S_n to the real axis, i.e. we have (50) for all $\varphi \in \mathbf{R}^1$. Thus we have shown the assertion of the theorem with $\varepsilon = \varepsilon_6$ in case 1.

Summarizing by starting with an initial condition in M_{ε_6} the Newton-Raphson method maps the set $U_{\varepsilon_6} \cap M_{\varepsilon_1}$ into itself and we have convergence towards the continuous function $S \in P$ satisfying the condition of invariance (1). Referring to (11) and (12) at the beginning of the proof of the theorem we are left to the case

2. $0 < a(S)(s_m) < 1$, $m = 0, 1, 2, \ldots, M-1$ (symmetry).

From the assumption that the derivative of S is continuous there follows from (10) and compactness that there exists a $C_a \in \mathbf{R}^1$ with $C_a > 0$ such that

$$a(S)(\varphi) > C_a > 0, \; \forall \varphi \in \mathbf{R}^1.$$

Following (8a) in Assumption 4.2 the circle map (2a) induced by S is invertible and we consider the transition from (3a) to (2.3.22) with (2.3.23). Additionally conditions (8b-g) in Assumption 4.2 are also valid for the inverse of the circle map (2a). Based on case 1 of the proof we conclude that the Newton-Raphson method is converging in an appropriate neighbourhood of the invariant curve for (2.3.22) with (2.3.23) and as solutions of (2.3.22) are invariant from (2.3.22) to (3a) the assertion of the theorem in case 2 is concluded.

◊

Example 4.4: As an example for (16a-e) we consider in (1.1.4) the angular map defined for $\varphi \in \mathbf{R}^1$

$$H(r, \varphi) = \varphi + r, r \in \mathbf{R}^1$$

with

$$r = v(\varphi) = \alpha \sin \varphi, \alpha \in \mathbf{R}^1, \ |\alpha| < 1$$

and the one-parametric set of circle maps

$$H(v(\varphi), \varphi) = \varphi + \alpha \sin \varphi.$$

H has fixed points

$$s_k = k \cdot \pi, k \in \mathbf{Z}$$

and is monotonically increasing with (16c). In (16) we have M = 2.

We note that this circle map is reflected in the first Newton-Raphson step of the example (6.2.1) with (6.2.2) in Section 6.2 (see (6.2.5)). For the following Newton-Raphson steps we have, however, to allow for general $v \in P$ with $v(s_k) = v(k\,\pi) = 0$, $k \in \mathbf{Z}$. In addition for the properties (16d), (16f) and (16g) we refer to Assumption 4.1.

◊

The following example illustrates the convergence proof in Theorem 4.15. The simplicity of the example is similar to Example 1.1. In Example 1.1 we illustrate the case without fixed point whereas here we have the case with fixed points. In both examples we find convergence in the first Newton-Raphson step.

Example 4.5: For $r \in \mathbf{R^1}$, $\varphi \in \mathbf{R^1}$ we consider in (1.1.4)

$$\begin{aligned} G(r, \varphi) &= \kappa \cdot r \\ H(r, \varphi) &= \varphi_0 \cdot \varphi \end{aligned} \tag{4.4.51}$$

where $\kappa \in \mathbf{R^1}$ and $\varphi_0 \in \mathbf{R^1}$ are given parameters. It is seen that the map (51) satisfy no periodicity condition as in (1.1.5). Furthermore it is decoupled in r and φ and obviously $S(\varphi) = 0$ is an invariant curve. We consider the initial condition $S_0(\varphi) = \lambda\, \varphi$ with $\lambda \in \mathbf{R^1}$. Reflecting Assumption 4.2 the circle maps $H(S_0(\varphi), \varphi)$ and $H(S(\varphi), \varphi)$ in (2a) and (3b) have a fixed point in $\varphi = 0$. Referring to (3) we have

$$h_0(\varphi) = \varphi_0 \cdot \varphi,$$

$$a_0(\varphi) = \kappa,$$

$$b_0(\varphi) = -\lambda \cdot \varphi_0 \cdot \varphi + \kappa \cdot r = -\lambda \cdot \varphi_0 \cdot \varphi + \lambda \cdot \kappa \cdot \varphi = \lambda\ \varphi \cdot (-\varphi_0 + \kappa),$$

$$d_0(\varphi_0 \cdot \varphi) - \kappa \cdot d_0(\varphi) = \lambda \cdot \varphi\ (-\varphi_0 + \kappa),\ \varphi \in \mathbf{R^1}.$$

The solution d_0 is

$$d_0(\varphi) = -\lambda\, \varphi$$

thus we have

$$S(\varphi) = S_1(\varphi) = S_0(\varphi) + d_0(\varphi) = 0.$$

In the context of iteration (3) we find using the series (4) for n = 0

$$d_0 = -\sum_{j=0}^{\infty} \frac{b_0 \circ h_0^j}{\prod_{i=0}^{j} a_0 \circ h_0^i}$$

and

$$b_0(h_0^n(\varphi)) = \lambda \cdot \varphi\ (\varphi_0)^n \cdot (-\varphi_0 + \kappa) = \lambda \cdot \varphi\ (\varphi_0)^{n+1} \left(\frac{\kappa}{\varphi_0} - 1 \right)$$

that

$$d_0(\varphi) = -\lambda\varphi\left(\frac{\kappa}{\varphi_0} - 1\right)\left(\left(\frac{\varphi_0}{\kappa}\right) + \left(\frac{\varphi_0}{\kappa}\right)^2 + \left(\frac{\varphi_0}{\kappa}\right)^3 + ...\right) = -\,\lambda\,\varphi.$$

◊

The theoretical investigation of the Newton-Raphson method presented in this thesis is concluded. We have shown convergence under the assumption that the circle map induced by the invariant curve has a finite and even number of fixed points and that these fixed points coincide with the fixed points of the circle map induced by the initial condition. Convergence proofs under more general assumptions are left to future research. However, it is seen that we have used the fixed point assumption extensively and our technique for showing convergence differs substantially from the convergence proofs in the literature (see e.g. Kantorovitch [10, 1981], Stoer [24, 1983]).

5. Algorithms for computing invariant curves

In this chapter we derive realizations of Algorithm 1. It is assumed that the linear functional equation (2.2.2a) discussed in Chapter 4 has a unique solution $f \in P$. We are concerned with the numerical computation of f. However, the evaluation of the series (2.3.6) and (2.3.7) described in Theorem 2.1 is not efficient because

- they only converge linearly although the Newton-Raphson method formulated in Algorithm 1 converges quadratically.

- they converge slowly if $|a(x)-1| < \varepsilon$, $x \in \mathbf{R}^1$ and $\varepsilon \to 0$.

- the evaluation of (2.3.7) requires the numerical inversion of h.

In Sections 5.1 and 5.2 we derive algorithms for approximating f that are not based on (2.3.6) and (2.3.7).

5.1. A spline method

Let

$$r = \{x_0, x_1, \ldots, x_{N-1}\} \tag{5.1.1}$$

be a partition of the interval $[0, 2\pi)$ with $x_0 = 0 < x_1 < x_2 < \ldots\ldots < x_{N-1} < 2\pi$. Then, the 2π-periodic continuation to the real axis is given by

$$x_{k+N} - x_k = 2\pi, \; k \in \mathbf{Z}. \tag{5.1.2}$$

Following [22, Schumaker] *B-splines* $M_{k,i}$ are recursively given by

$$M_{k,i}(x) = \frac{x - x_k}{x_{k+i} - x_k} M_{k,i-1}(x) + \frac{x_{k+i} - x}{x_{k+i} - x_k} M_{k+1,i-1}(x) \tag{5.1.3}$$

for $i = 2, 3, \ldots, k \in \mathbf{Z}, x \in \mathbf{R}^1$

where

$$M_{k+1,1}(x) = \begin{cases} \dfrac{1}{x_{k+1} - x_k}, x \in [x_k, x_{k+1}) \\ \\ 0, \qquad\qquad \text{else} \end{cases} . \tag{5.1.4}$$

With

$$B_{k,i}(x) = (x_{k+i} - x_k)\, M_{k,i}\,(x) \tag{5.1.5}$$

we have

$$\sum_{k \in Z} B_{k,i}(x) = 1,\ i \in \mathbf{N}$$

[22, Schumaker]. In the following we work with the *normed B-splines* defined by (5). In this section and in Chapter 7 and 8 we use (3), (4) and (5) with i = 4 and in order to simplify the notation we suppress the index 4. However, the algorithm derived in this section can easily be formulated for any $i \in \mathbf{N}$. We have

$$B_k(x) = 0,\ x \notin (x_k, x_{k+4}) \tag{5.1.6}$$

and B_k is a cubic polynomial in each interval $[x_{k+m}, x_{k+m+1}]$, $0 \le m \le 3$. In addition, B_k is 2-times differentiable in the knots x_{k+m}, $0 \le m \le 4$.

In the following we derive an algorithm for approximating the unique solution $f \in P$ of (2.3.2) described in Theorem 2.1. We start by expressing the function values $f(x_j)$, j = 0,…., N – 1 as a linear combination of the B-splines:

$$f(x_j) = \sum_{k=0}^{N-1} \hat{f}_k B_k(x_j),\ j = 0,\ldots., N-1 \tag{5.1.7}$$

where the coefficients $\hat{f}_k \in \mathbf{R}^1$, $0 \le k \le N - 1$ are unknown. By using the linearity of the operator L defined in (2.3.3) we have

$$Lf(x_j) = \sum_{k=0}^{N-1} \hat{f}_k L B_k(x_j),\ j = 0,\ldots., N-1. \tag{5.1.8}$$

We consider the expansions

$$LB_k(x_j) = \sum_{m=0}^{N-1} a_{k,m} B_m(x_j),\ k = 0,\ldots, N-1,\ j = 0,\ldots, N-1, \qquad (5.1.9a)$$

$$b(x_j) = \sum_{k=0}^{N-1} b_k B_k(x_j),\ j = 0,\ldots, N-1 \qquad (5.1.9b)$$

where the coefficients $a_{k,m} \in \mathbf{R}^1$, $0 \le m, k \le N-1$, $b_k \in \mathbf{R}^1$, $0 \le k \le N-1$ are unknown. Substituting (9a) in (8) and equating coefficients for B_k, $0 \le k \le N-1$ in $Lf(x_j) = b_j$, $j = 0,\ldots, N-1$ yield the system of linear equations

$$A\hat{f} = b,\ A = \begin{bmatrix} a_{0,0} & \cdot & \cdot & a_{0,N-1} \\ \cdot & & & \cdot \\ & & & \\ \cdot & & & \cdot \\ a_{N-1,0} & \cdot & \cdot & a_{N-1,N-1} \end{bmatrix},\quad b = \begin{bmatrix} b_0 \\ \cdot \\ \\ \cdot \\ b_{N-1} \end{bmatrix} \qquad (5.1.10)$$

for the unknown vector $\hat{f} = (\hat{f}_0,\ldots,\hat{f}_{N-1})$.

In the following we outline the computation of $f(x_j)$, $j = 0,\ldots, N-1$ in (7). 3 steps have to be considered.

In *step 1* we compute the elements of A and b in (10). Assuming that the coefficient functions $h(x) = x + p(x)$, $p \in P$, $a \in P$, $b \in P$ of (2.3.2) are given, the left side of (9a) is evaluated by using (1), (2), (3), (4) and

$$LB_k(x) = B_k(h(x)) - a(x) \cdot B_k(x),\ x \in \mathbf{R}^1,\ k = 0,\ldots, N-1.$$

From (2), (6) and (9a) it follows that the computation of A and b be requires the solution of the N systems of linear equation

$$TA = (LB_0,\ldots,LB_{N-1}),$$

and the system

$$T(b_0,\ldots,b_{N-1}) = (b(x_0),\ldots,b(x_{N-1}))$$

where T is given by the N-dimensional matrix

$$\begin{bmatrix} B_{-2}(x_0) & B_{-1}(x_0) & 0 & & 0 & B_{N-3}(x_0) \\ B_{-2}(x_1) & B_{-1}(x_1) & B_0(x_1) & & 0 & 0 \\ 0 & B_{-1}(x_2) & B_0(x_2) & & & \\ & & & & & \\ & & & B_{N-5}(x_{N-3}) & B_{N-4}(x_{N-3}) & 0 \\ 0 & 0 & & B_{N-5}(x_{N-2}) & B_{N-4}(x_{N-2}) & B_{N-3}(x_{N-2}) \\ B_{-2}(x_{N-1}) & 0 & & 0 & B_{N-4}(x_{N-1}) & B_{N-3}(x_{N-1}) \end{bmatrix}. \quad (5.1.11)$$

T describes the transformation between the coefficient in the basis of the B-splines and function values. In the special case of equidistant knots we find with (4) and (11)

$$T = \frac{1}{6}\begin{bmatrix} 4 & 1 & 0 & & & & & 0 & 1 \\ 1 & 4 & 1 & & & & & 0 & 0 \\ 0 & 1 & 4 & 1 & & & & & \\ & & & & & & & & \\ & & & & 1 & 4 & 1 & & \\ 0 & 0 & & & & 1 & 4 & 1 \\ 1 & 0 & & & & & 1 & 4 \end{bmatrix}.$$

From the special structure of T it follows that the LU decomposition requires O(N) floating point operations. Together with the N+1 forward and backward substitutions the computation of A and b needs $O(N^2)$ floating point operations. The numerical experiments show that, in general, A contains no exact zeros.

In step 2 we solve (10). Mainly direct methods are used. However, for large systems (10) two iterative solvers are also considered (see Sections 5.3 and 6.3).

In step 3 the approximation $f = (f_0,\ldots,f_{N-1})$ of $f(x_j)$, $j = 0,\ldots, N-1$ is found by evaluating

$$f = T \cdot \hat{f}. \quad (5.1.12)$$

The algorithm below combines Algorithm 1 with the considerations of this section. The n^{th} Newton-Raphson approximation S_n is computed with varying partitions (1). If a knot in the $(n + 1)^{st}$ Newton-Raphson step is not considered in the n^{th} Newton-Raphson step we interpolate by using a 2π-

periodic cubic periodic spline function that represents S_n. We solve (10) by increasing the dimension N of the system of linear equations. S_{n+1} is approximated in partitions with increasing number of knots. The partitions are chosen such that the grid of the next higher dimension contains the knots of the previous partition. The accuracy of the computation is checked in the maximum norm given for a vector $v = (v_0,\ldots, v_{N-1})$ in $\mathbf{R}^N$ by

$$\|v\|_{\max} = \max_{0\le k\le N-1} |v_k|. \tag{5.1.13}$$

Algorithm 2 (spline method)

Notations:
1. S_n: Newton-Raphson approximation in the n^{th} Newton-Raphson step;
2. r_N, t_M: Vectors of length N, M, respectively containing partitions (1) of the interval $[0, 2\pi)$;
3. $\hat{S}_n$: Vector containing approximations of S_n in $x_j \in r$;
4. $\hat{S}_{n+1}$: Vectors containing approximation of S_{n+1} in $x_j \in r_N$;
5. d_n: Vectors containing approximations of the Newton-Raphson correction d_n in $x_j \in r_N$.

Given:
1. r: Partition (1) of the interval $[0, 2\pi)$;
2. $\hat{S}_0$: Vector containing function values of an initial approximation S_0 in $x_j \in r$;
3. ε: requested precison of the computed invariant curve;
4: $N_n^{\max}$: Maximal number of considered knots in the n^{th} Newton-Raphson step.

BEGIN

$n = 0$;

REPEAT

$d_n = 0$; $\hat{S}_{n+1} = 0$; $M = 2$;

The Newton-Raphson approximation d_n is represented by a periodic cubic spline and the derivative in (2.2.2c) is computed by interpolating the derivative of this spline function.

REPEAT

$t_M = r_N$;

1. Choose partition r_N with $r_N \supset t_M$, $N > M$;

2. Compute the LU decomposition of T given by (11). Evaluate the left side of (9) and compute the elements of A and b by backward and forward substitution;

3. Solve the system of linear equations A $\hat{d}_n$ = b for the vector $\hat{d}_n$ (see (10));

4. Evaluate $d_n = T\,\hat{d}_n$ (see (12));

5. $\tilde{S}_{n+1} = \hat{S}_n + d_n$, $x_j \in r_N$;

6. Let $\tilde{S}_{n+1}$ be the restriction of $\hat{S}_{n+1}$ to the partition t_M;

UNTIL $\left\| \tilde{S}_{n+1} - \hat{S}_{n+1} \right\|_{max} < \varepsilon$ or $N = N_n^{max}$;

$\hat{S}_n = \hat{S}_{n+1}$; $n = n+1$;

UNTIL $\left\| d_n \right\|_{max} < \sqrt{\varepsilon}$;

END.

◊

In Chapter 6 and 7 we describe numerical experiments that are based on Algorithm 2. In most cases equidistant knots are used. However, considering an invariant curves with low differentiability, it is illustrated in Section 6.3 that the accuracy is improved by introducing nonequidistant knots.

5.2. A Fourier Method

In Sections 4.2 and 4.3 we discussed a unique solution $f \in P$ of (2.2.2a) discussed in Chapter 4. Fourier series are a natural choice for approximating f as the 2π-periodicity is inherent.

We start by deriving an infinite system of linear equations for the complex Fourier coefficients of the solution f in (4.2.1). Thus we consider

$$f(x) = \sum_{k=-\infty}^{\infty} f_k e^{ikx} \tag{5.2.1}$$

where the Fourier coefficient $f_k \in \mathbf{C}$, $k \in \mathbf{Z}$ are unknown. As f is real-valued we have

$$f_k = \overline{f_{-k}}\,,\ k \in \mathbf{Z}.$$

Following (2.3.3) we consider the linear map

$$L(f) = f \circ h - a \cdot f \tag{5.2.2}$$

with

$$L(f) = b.$$

Substituting (1) in (4.2.1), (4.3.1), respectively and using (2.3.1) with $h(x) = x + p(x)$, $p(x) \in P$ yield

$$\sum_{k=-\infty}^{\infty} f_k e^{ikx} e^{ikp(x)} - a(x) \sum_{k=-\infty}^{\infty} f_k e^{ikx} = b(x),$$

thus

$$\sum_{k=-\infty}^{\infty} f_k e^{ikx} (e^{ikp(x)} - a(x)) = b(x).$$

It is assumed that the Fourier series

$$e^{ikx} (e^{ikp(x)} - a(x)) = \sum_{m=-\infty}^{\infty} a_{m,k} e^{imx}\,,\ k \in \mathbf{Z}$$

with

$$a_{m,k} = \frac{1}{2\pi}\int_0^{2\pi}(e^{ikp(x)} - a(x))e^{ikx}e^{-imx}dx\,,\ m \in \mathbf{Z},\ k \in \mathbf{Z}, \qquad (5.2.3a)$$

and

$$b(x) = \sum_{k=-\infty}^{\infty} b_k e^{ikx}$$

with

$$b_k = \frac{1}{2\pi}\int_0^{2\pi} b(x)\, e^{-ikx}dx\,,\ k \in \mathbf{Z} \qquad (5.2.3b)$$

exist. With (4.2.1) we obtain

$$\sum_{k=-\infty}^{\infty} f_k \sum_{m=-\infty}^{\infty} a_{m,k}e^{imx} = \sum_{k=-\infty}^{\infty} b_k e^{ikx}\,.$$

By equating coefficient for e^{ikx}, $k \in \mathbf{Z}$, we find the infinite system of linear equations

$$\begin{bmatrix} & & & \cdot & & & \\ & & & \cdot & & & \\ & & a_{-1,-1} & a_{-1,0} & a_{-1,1} & & \\ \cdot & \cdot & a_{0,-1} & a_{0,0} & a_{0,1} & \cdot & \cdot \\ & & a_{1,-1} & a_{1,0} & a_{1,1} & & \\ & & & \cdot & & & \\ & & & \cdot & & & \end{bmatrix} \begin{bmatrix} \cdot \\ \cdot \\ f_1 \\ f_0 \\ f_{-1} \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ b_1 \\ b_0 \\ b_{-1} \\ \cdot \\ \cdot \end{bmatrix} \qquad (5.2.4)$$

for the unknown coefficients $f_k \in \mathbf{C}$, $k \in \mathbf{Z}$. From the assumptions $h(x) \in \mathbf{R^1}$, $a(x) \in \mathbf{R^1}$ and $b(x)$, $x \in \mathbf{R^1}$ it follows

$$a_{m,k} = \overline{a_{-m,-k}}\,,\ m, k \in \mathbf{Z},$$

$$b_k = \overline{b_{-k}}\,,\ k \in \mathbf{Z}$$

in (3).

Example 5.1: Let

$$h(x) = x + h_0 + h_1 \sin x,\ h_0 \in \mathbf{R^1},\ h_1 \in \mathbf{R^1},$$

$$a(x) = a_0 + a_l \sin x,\ a_0 \in \mathbf{R^1},\ a_l \in \mathbf{R^1}, \tag{5.2.5}$$

$$b(x) = 1$$

in (4.2.1), (4.3.1) respectively. (3a) yields

$$a_{m,k} = \frac{1}{2\pi}\int_0^{2\pi}(e^{ikh_0}e^{ikh_1 \sin x} - a(x))\, e^{ikx}e^{-imx}\, dx,\ m \in \mathbf{Z},\ k \in \mathbf{Z}. \tag{5.2.6}$$

By considering the Bessel functions $J_m(y)$, $y \in \mathbf{R^1}$, m = 0, 1, 2,…. we have

$$e^{iy\sin x} = \sum_{m=-\infty}^{\infty} J_m(y)e^{imx},\ y \in \mathbf{R^1}$$

[1, Abramowitz]. With $y = kh_l$ and (6) we have:

$$a_{m,k} = e^{ikh_0} J_{m-k}(kh_1) - \frac{1}{2\pi}\int_0^{2\pi} a(x)\, e^{ikx}e^{-imx}\, dx,\ m \in \mathbf{Z},\ k \in \mathbf{Z}.$$

The first term of the difference yields the infinite matrix

$$\begin{bmatrix} & & & \cdot & & & \\ & & & \cdot & & & \\ & & e^{-ih_0}J_0(-h_1) & 0 & e^{ih_0}J_{-2}(h_1) & & \\ \cdot & \cdot & e^{-ih_0}J_1(-h_1) & 1 & e^{ih_0}J_{-1}(h_1) & \cdot & \cdot \\ & & e^{-ih_0}J_2(-h_1) & 0 & e^{ih_0}J_0(h_1) & & \\ & & & \cdot & & & \\ & & & \cdot & & & \end{bmatrix}.$$

Based on the asymptotic behaviour of the Bessel functions [1] the structure of this matrix can be described. Assuming $h_1 \neq 0$, the matrix contains no exact zeros with the exception of $a_{m,0}$, $m \in \mathbf{Z}$, $m \neq 0$. However, the elements in the upper right corner and the lower left comer are small.

The second term of this difference is an infinite tridiagonal matrix

$$\begin{bmatrix} a_0 & a_1 & 0 & & & 0 \\ a_1 & a_0 & a_1 & 0 & & \\ 0 & a_1 & a_0 & a_1 & & \\ & 0 & a_1 & a_0 & & \\ & & 0 & & & \\ & & & & & \\ 0 & & & & & \end{bmatrix}.$$

The main diagonal contains $a_0 \in \mathbf{R}^1$ and in the upper and the lower diagonal we have $a_l \in \mathbf{R}^1$.

◊

From Example 5.1 we see that, although assuming only 2 nontrivial Fourier coefficient in (5) for the coefficients functions a and h in (4.2.1), the matrix $A = (a_{m,k})$, $m \in \mathbf{Z}$, $k \in \mathbf{Z}$ contains no exact zeros. The numerical experiments show, however, that the linear systems that has to be solved for computing the solution f of (4.2.1) differ in most case little from the system discussed in Example 5.1.

In the following we derive an algorithm for computing the solution f of (2.3.2). Real Fourier series are preferred to complex Fourier series because

a. Only real-valued functions are considered.

b. The implementation does not request complex arithmetic, i.e. we do not generate additional operations by the complex arithmetic.

c. As in Section 5.1 we only have to solve linear equations with a real coefficient matrix.

d. Infinity of the linear system is only approached in two and not in four directions.

Let

$$f(x) = \sum_{k=0}^{\infty} f_k B_k(x)\,,\ x \in \mathbf{R^1}$$

where

$$B_k(x) = \begin{cases} \cos(\frac{k}{2}x) & k = 0, 2, 4, \ldots.. \\ \sin(\frac{k+1}{2}x) & k = 1, 3, 5, \ldots.. \end{cases} \qquad (5.2.7)$$

and $f_k \in \mathbf{R^1}$, k = 0, 1, 2,…. and the linearity of the operator L defined in (2) yields:

$$Lf(x) = \sum_{k=0}^{\infty} f_k LB_k(x) = b(x). \qquad (5.2.8)$$

It is assumed that

$$LB_k(x) = (B_k \circ h)(x) - a(x)\, B_k(x),\ k = 0, 1, 2, \ldots. \qquad (5.2.9)$$

and b can be represented by real Fourier series, i.e. there exists coefficients $a_{m,k}$, m = 0, 1, 2,…., k = 0, 1, 2,…. and b_k, k = 0, 1, 2,…. such that

$$LB_k(x) = \sum_{m=0}^{\infty} a_{m,k} B_m(x)\,,\ k = 0, 1, 2, \ldots.,\ b(x) = \sum_{k=0}^{\infty} b_k B_k(x)\,, \qquad (5.2.10)$$

where

$$a_{0,k} = \frac{1}{2\pi}\int_0^{2\pi} LB_k(x)dx\,,\ k = 0, 1, 2, \ldots.,$$

$$a_{m,k} = \frac{1}{\pi}\int_0^{2\pi} LB_k(x)B_m(x)\,dx\,,\ k = 0, 1, 2, \ldots.,\ m = 1, 2, \ldots.,$$

$$b_0 = \frac{1}{2\pi}\int_0^{2\pi} b(x)\,dx\,,$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} b(x)B_k(x)\,dx\,,\ k = 1, 2, \ldots..$$

With (8), (9) and (10) we have:

$$\sum_{k=0}^{\infty} f_k \sum_{m=0}^{\infty} a_{m,k} B_m(x) = \sum_{k=0}^{\infty} b_k B_k(x).$$

Equating coefficient for B_k, k = 0, 1, 2,…. yields the infinite system of linear equations

$$\begin{bmatrix} a_{0,0} & a_{0,1} & a_{0,2} & . & . & . \\ a_{1,0} & a_{1,1} & a_{1,2} & . & . & . \\ a_{2,0} & a_{2,1} & & & & \\ . & & & & & \\ . & & & & & \\ . & & & & & \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \\ \\ \\ \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \\ \\ \\ \end{bmatrix} \qquad (5.2.11)$$

for the unknown coefficients $f_k \in \mathbf{R}^1$, k = 0, 1, 2,….. Finite subsystems of order N are considered. In order to apply the Fast Fourier Transform (F.F.T.) we choose $N = 2^i$, i = 0, 1, 2,…. and evaluate the right side of (9) for k = 0,…., N – 1 and the function b in

$$x_j = \frac{2\pi}{N},\ j = 0, 1,\ldots,N-1. \qquad (5.2.12)$$

By applying F.F.T. we obtain the finite system of linear equation

$$\hat{A}\,\hat{f} = \hat{b}, \hat{A} = \begin{bmatrix} \hat{a}_{0,0} & . & . & \hat{a}_{0,N-1} \\ . & & & . \\ & & & \\ . & & & . \\ \hat{a}_{N-1,0} & . & . & \hat{a}_{N-1,N-1} \end{bmatrix}, \quad \hat{b} = \begin{bmatrix} \hat{b}_0 \\ . \\ \\ . \\ \hat{b}_{N-1} \end{bmatrix} \qquad (5.2.13)$$

for the unknown vector $\hat{f} = (\hat{f}_0,\ldots,\hat{f}_{N-1})$. The components of $\hat{f}$ are approximations of the first N Fourier coefficients of the solution f of (4.2.1). The finite system (13) is considered as an approximation of the infinite system (11).

The following Algorithm 3 combines Algorithm 1 (see Section 2.2) with the considerations of this section. Assuming that the Fourier coefficients of an initial approximation S_0 are given, it computes the Fourier coefficients of an invariant curve S. Essentially the algorithm consists of two steps:

In step 1 the linear sub problem (2.2.2a) is solved by increasing the dimension N in (13). We approximate the Fourier coefficients of the n^{th} Newton-Raphson correction d_n in Algorithm 1.

In step 2 The Fourier coefficients computed in step 1 are added to the approximation of the Fourier coefficients of S_n, n = 0, 1, 2,….. More details of the implementation are given below.

Algorithm 3 (Fourier method)

Notation: 1. s_n: Approximation of the Fourier coefficients in the n^{th} Newton-Raphson step, n = 0, 1, 2,…. The length of the vector can increase during the computation of the invariant curve;

2. $\delta_n, \hat{\delta}_n$: Vector of length N, M, respectively, containing the first N, M, respectively Fourier coefficient of the Newton-Raphson correction, d_n, n = 0, 1, 2,…..

Given: 1. s_0: Fourier coefficients of the initial approximation S_0;

2. ε: Requested precision of the computed invariant curve.

BEGIN

$n = 0; s_0 = 0;$

REPEAT

$N = 1; \delta_n = 0;$

REPEAT

$M = N;\ \hat{\delta}_n = \delta_n;\ N = 2N;$

1. Evaluate the function h_n, a_n, b_n defined by (2.2.2b)-(2.2.2d) in the knots (13). It is seen from (2.2.2c) and (2.2.2d) that the evaluation of a_n and b_n requires the evaluation S_n' and S_n in $H(S_n(x), x)$. In general $H(S_n(x), x)$ is not a knot considered in (13). We interpolate S_n and its derivatives S_n' by considering the Fourier polynomial and its derivative for evaluating S_n and S_n' in $H(S_n(x), x)$. We use the algorithm of Reinsch (see Stoer [24, 1983]);

2. Compute the matrix A and the vector b of the system of linear equation (13) by applying F.F.T. to
   $LB_k(x_j) = (B_k \circ h_n)(x_j) - a_n(x_j)\, B_k(x_j)$, $0 \le k \le N-1$ and
   $b_n(x_j)$, $0 \le j \le N-1$,
   where B_k is given by (7) and x_j is given by (12);

3. Solve the system of linear equation $A\ \delta_n = b$ (see (13)) for the vector δ_n and add M zeros to the components of $\hat{\delta}_n$ in order to get a vector of length N;

UNTIL $\| \hat{\delta}_n - \delta_n \|_{max} < \varepsilon$;

$s_{n+1} = s_n + \delta_n$; $n = n + 1$;

(If s_n and δ_n have different lengths we extend the vector with shorter length by zeros such that both vectors have the same length.)

UNTIL $\| \delta_n \|_{max} < \sqrt{\varepsilon}$;

END.

◊

Chapters 6, 8 and 9 report on numerical experiments that are based on Algorithm 3.

5.3. Iterative solvers of systems of linear equations

Following step 3 of Algorithm 2 (see Section 5.1) and step 3 of Algorithm 3 (see Section 5.2) systems of linear equations

$$A\,x = b,\ A \in \mathbf{R}^{\mathbf{NxN}},\ b \in \mathbf{R}^{\mathbf{N}} \qquad (5.3.1)$$

for the unknown vector $x \in \mathbf{R}^{\mathbf{N}}$ are considered in this section. Applying direct methods to (1), we encounter no numerical problems for solving (1), i.e. the accuracy of the solution $x \in \mathbf{R}^{\mathbf{N}}$ is close to machine precision. As described in Example 5.1 (see Section 5.2) the matrices arising in this thesis have in general no exact zeros. Thus, solving (1) requires $O(N^3)$ floating point operations.

Iterative solvers are based on the operation matrix times vector, i.e. for one iteration $O(N^2)$ floating point operations are needed. It follows that they are more efficient if convergence is achieved in much less than N iterations. In particular, fast convergence is expected if the required accuracy in Algorithm 1 is low. In the following we present the iterative solvers considered in Section 6.3, paragraph C.

For solving (1), we introduce the vector iteration

$$x_{k+1} = x_k - A_0^{-1} \cdot (A\,x_k - b),\ k = 0, 1, 2, \ldots. \qquad (5.3.2)$$

with the initial condition $x_0 \in \mathbf{R}^{\mathbf{N}}$ and it is assumed that $A_0 \in \mathbf{R}^{\mathbf{NxN}}$ is nonsingular. Let $\Lambda(C)$ denote the *spectral radius* of a matrix $C \in \mathbf{R}^{\mathbf{NxN}}$ [24, Stoer]. From $\Lambda\,(I - A_0^{-1}A) < 1$ it follows that the sequence x_k, $k = 0, 1, 2, \ldots.$ converges towards the solution x of (1). The convergence is fast if $\Lambda(I - A_0^{-1}A) << 1$ [24]. From $\Lambda(I - A_0^{-1}A) = 1 - \Lambda\,(A_0^{-1}\,A)$ it is concluded that a good approximation A_0^{-1} of A^{-1} is required. However, the main idea of the numerical realisation of (2) is that A_0^{-1} is much easier to compute than A^{-1}.

Following the description of A in Example 5.1 we choose for A_0 a band matrix with $q = 0, 1, 2, \ldots.$ diagonals below and beyond the main diagonal. $q = 2$ is illustrated in Figure 2. The rest of the elements are set to zero. In order to generate x_k, $k = 1, 2, \ldots.$ the LU decomposition of A_0 has only to be computed once. For matrices A_0 considered in this thesis the numerical experiments show that the pivots for the LU decomposition can be chosen in the main diagonal of A_0. The LU decomposition preserves the band structure of A_0, i.e. the LU decomposition can be done on the nontrivial elements of A_0.

$A_0 =$

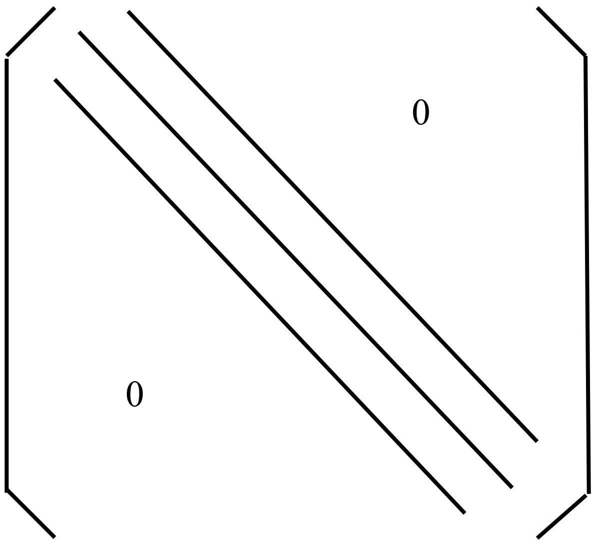

Figure 2

Using the max norm, defined by (5.1.13) the implementation of (2) yields the following algorithm:

Algorithm 4 (Computation of the solution x $\in$ $\mathbf{R}^N$ in (1) with iteration (2))

Given: 1. System of linear equation $Ax = b$;
2. Matrix A_0;
3. Initial approximation x_0;
4. Tolerance ε;

BEGIN

$k = -1$; Compute the LU decomposition of A_0;

REPEAT

$k = k + 1$;

1. Compute the vector $b = A\, x_k - b$;
2. Solve the system $A_0 u = b$ for $u \in \mathbf{R}^N$ by forward and backward substitution;
3. Let $x_{k+1} = x_k - u$;

UNTIL $\|x_{k+1} - x_k\|_{max} < \varepsilon$;

END;

◊

By using the vector iteration (2) the following difficulties arise:

1. In order to ensure convergence of Algorithm 4 the spectral radius $\Lambda(A_0^{-1}A)$ is required to be known.
2. We cannot determine a priori the optimal choice of q in A_0.

If A is symmetric positive definite *the conjugate residual method* and *the conjugate residual method* avoid the difficulties mentioned above. They converge in N steps at the most, independent of the spectral radius A of A. In addition, no approximation of A^{-1} is needed [23, Stiefel]. However, as describe in Example 5.1, the matrices arising in this thesis are in general not symmetric. In [5, Eisenstat] a generalisation of conjugate residual method to nonsymmetric matrices is discussed. The algorithm is presented below.

Algorithm 5 (Computation of the solution $x \in R^N$ in (1) with the generalised conjugate residual method)

Given: 1. System of linear equation $Ax = b$;
2. Initial approximation x_0 with residual $r_0 = Ax_0 - b$;
3. Tolerance ε;

BEGIN

$p_0 = r_0$; $k = -1$;

REPEAT

$k = k + 1$;

1. Let

$$x_{k+1} = x_k + \alpha_k p_k,$$
$$r_{k+1} = r_k - \alpha_k Ap_k$$

where

$$\alpha_k = \frac{(r_k, Ap_k)}{(Ar_k, Ap_k)};$$

2. Compute

$$p_{k+1} = r_{k+1} + \sum_{i=0}^{k} \beta_i^k p_i \tag{5.3.3}$$

where

$$\beta_i^k = \frac{(Ar_{k+1}, Ap_i)}{(Ap_i, Ap_i)}, \; 0 \le i \le k;$$

UNTIL $\|x_{k+1} - x_k\|_{max} < \varepsilon$;

END;

◊

We proceed by discussing the above algorithm. With (3) it is concluded:

$$(Ap_i, Ap_j) = 0, \; i \ne j, \tag{5.3.4}$$

i.e. the vectors p_j, j = 2, 3,.... are A^TA-orthogonal. Based on step 1 of Algorithm 5 we have

$$r_{k+1} = r_k - \frac{(r_k, Ap_k)}{(Ap_k, Ap_k)} Ap_k.$$

By using induction on k it is shown that

$$(r_{k+1}, Ap_j) = 0, \; 0 \le j \le k, \; k = 0, 1, 2, \ldots. \tag{5.3.5}$$

From $Ap_j \in \mathbf{R}^N$, (4) and (5) it follows $r_{N+1} = 0$, i.e. the computation of $x \in \mathbf{R}^N$ in (1) requires at most N iterations independent of the spectral radius Λ of A and the tolerance ε. In the special case of A symmetric and positive definite the properties (4) and (5) are preserved by considering the simplified version

$$p_{k+1} = r_{k+1} + \beta_k p_k$$

where

$$\beta_k = \frac{(Ar_{k+1}, Ap_k)}{(Ap_k, Ap_k)}, \; k = 0, 1, 2, \ldots.$$

of Algorithm 5 [5, Eisenstat]. Algorithm 5 reduces to *the conjugate residual method* [23, Stiefel].

We apply Algorithm 5 to the linear systems arising in Algorithm 2 and 3. Our numerical experiments show that convergence is observed if we replace in (3)

$$\sum_{i=0}^{k} \beta_i^k p_i$$

by

$$\sum_{i=k-m}^{k} \beta_i^k p_i \,.$$

We conjecture that the choice m $\approx$ 5 is successful is most cases.

Algorithms 2 and 3 show that the matrices A in (1) are not explicitly given. They need to be generated numerically. As discussed in Sections 5.1 and 5.2 we need $O(N^2)$ floating point operations for generating the matrix A in (1). As Gauss elimination requires $O(N^3)$ operations it is concluded that the iterative solvers formulated in Algorithms 4 and 5 are more efficient if they converge in much less than N iterations. An analysis of Algorithms 4 and 5 in terms of computer time, number of iterations etc. is given in Section 6.3.

6. Explicitly given maps

6.1. A test example

In this section we expand on Example 1.3 considered in the introduction (Section 1.4) of this thesis. The application of Algorithm 1 and 3 to the map (1.4.8) with (1.4.9) is described. We start by discussing the 1st Newton-Raphson step. With n = 0 and

$$S_0(\varphi) = 0 \tag{6.1.1}$$

the evaluation of (2.2.2b-d) yields

$$h_0(\varphi) = \varphi + 1,$$

$$a_0(\varphi) = 1.5,$$

$$b_0(\varphi) = 0.1 + \sin \varphi.$$

Let

$$d_0(\varphi) = \sum_{k=0}^{\infty} \delta_k B_k(\varphi),$$

where B_k, k = 0, 1, 2,…. is given by (5.2.7) and $\delta_k \in \mathbf{R}^1$, k = 0, 1, 2,…. is unknown. By using comparison of coefficients for B_k we find for (5.2.11) with $f(\varphi) = d_0(\varphi)$ in (5.2.8) the infinite linear system of linear equations

$$\begin{bmatrix} -0.5 & 0 & 0 & 0 & 0 & \\ 0 & -1.5+\cos 1 & -\sin 1 & 0 & 0 & \\ 0 & \sin 1 & -1.5+\cos 1 & 0 & 0 & \\ 0 & 0 & 0 & -1.5+\cos 2 & -\sin 2 & \\ 0 & 0 & 0 & \sin 2 & -1.5+\cos 2 & \\ 0 & 0 & 0 & 0 & 0 & -1.5+\cos 3 \\ . & . & . & . & . & \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ . \\ \\ \\ \end{bmatrix} = \begin{bmatrix} 0.1 \\ 1 \\ 0 \\ . \\ \\ \\ \end{bmatrix}. \tag{6.1.2}$$

The solution of the subsystem of order 3 is

$$d_0(\varphi) = \delta_0 + \delta_1 \sin \varphi + \delta_2 \cos \varphi$$

where

$$\delta_0 = -0.2,$$

$$\delta_1 = -0.58909'9361, \tag{6.1.3}$$

$$\delta_2 = -0.51652'7259$$

and with

$$\delta_k = 0, \ k \geq 3$$

we find a solution of the infinite system of linear equations (2). It is seen that

$$\tau = 1, \ \kappa = 1.5 \text{ and } T(\varphi) = 0.1 + \sin\varphi \tag{6.1.4}$$

is a set of parameters considered in map (1.4.1) (Example 1.2) and the condition (1.1.7) of invariance of the map (1.4.1) with (4) is identical to the first Newton-Raphson step (i.e. (2.2.2a) with n = 0) for the map (1.4.8) with (1.4.9). Consequently there follows that, as discussed in Example 1.2, we find with (3) the unique real-valued 2π-periodic trigonometric polynomial for the invariant curve S of (1.4.1) and furthermore we have that the solution (3) of (2) is unique. Applying Algorithm 1 to the map (1.4.8) by using (1) we have

$$S_l(\varphi) = d_0(\varphi).$$

Summarizing it is shown that the Fourier coefficients of the 1st Newton-Raphson step are explicitly known.

The Newton-Raphson approximations S_2, S_3, S_4,…. cannot be calculated explicitly. Starting with (1) we use Algorithm 3 for computing the invariant map given by (1.4.8) with (1.4.9) for the tolerance $\varepsilon = 10^{-5}$, $\varepsilon = 10^{-10}$, respectively. The Newton-Raphson method converges in 3, 4, respectively steps, i.e. Algorithm 3 goes 3, 4, respectively times through the outer loop. In the inner loop the Fourier coefficients of the Newton-Raphson correction are approximated by solving (5.2.13). We summarize the results of the computation:

1. In Table 1 the norms of the difference between successive solutions of the linear system (5.2.13) for $N = 2^i$, $3 \le i \le 6$ are shown, i.e. the values $\left\| d_n - \tilde{d}_n \right\|_{max}$ in Algorithm 3 are given. Table 1 indicates that by doubling the dimension of the system (5.2.13) the number of correct digits is doubled too.

n \ N	8	16	32	64
0	0.00000'0000	0.00000'0000		
1	0.10714'6040	0.00099'7250	0.00008'7681	0.00000'00021
2	0.09260'6429	0.00099'7250	0.00012'3658	0.00000'00021
3	0.09255'7006	0.00165'4441	0.00000'6876	0.00000'00021
4	0.09257'2675	0.00165'3940	0.00000'6875	0.00000'00001

Table 1

2. We observe that the first Fourier coefficients are converging faster than $\left\| d_n - \tilde{d}_n \right\|_{max}$. Figure 3 illustrates the behaviour of the Fourier series. In Table 2 we show the 3^{rd} component of the vector d_n (coefficient of $\cos \varphi$).

n \ N	8	16	32	64
0	-0.51652'7259	-0.51652'7259		
1	-0.65668'0784	-0.65990'3481	-0.65994'7431	-0.65994'7428
2	-0.68365'9820	-0.68606'4661	-0.68621'4557	-0.68621'4182
3	-0.68671'9854	-0.68658'6740	-0.68658'1431	-0.68658'1507
4	-0.68658'1669	-0.68658'1584	-0.68658'1583	-0.68658'1583

Table 2

3. Based on remark 2 above, we compute the invariant curve of Example 1.3 with Algorithm 3 by replacing ε by $\sqrt{\frac{\varepsilon}{10}}$ in the inner loop. With 5, 10, respectively digits tolerance the Newton-Raphson method converges in 3, 4, respectively steps. The inner loop needs systems of linear equations (5.2.13) with N = 64, N = 128, respectively for convergence. The algorithm requires 104, 264, respectively calls of the maps T and H given by (1.4.8) and (1.4.9).

Following the description of Example 1.3 in Chapter 1 the computation of S by implementing the fixed point iteration (1.4.7) needs considerably more calls of T and H. The comparison of the numerical values for the invariant curve obtained by the iteration (1.4.7) with the computation using the Newton-Raphson method yields a test of our implementation.

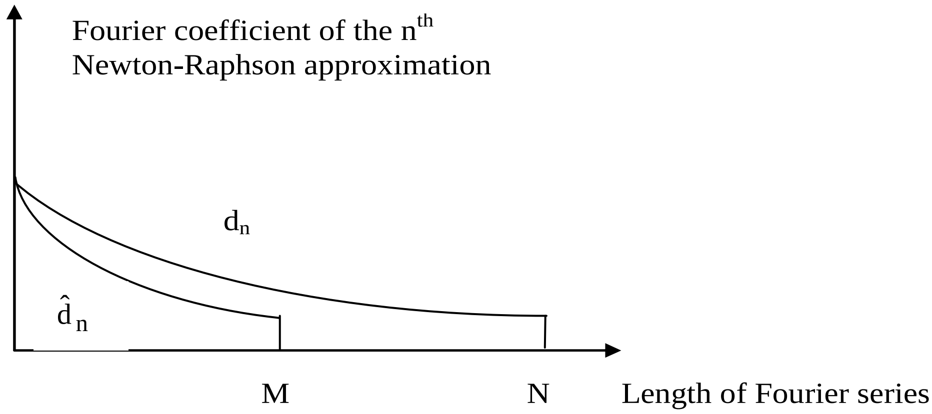

Figure 3

6.2. A map with two invariant curves

In this section and Section 6.3 we consider the realization

$$G(r, \varphi) = p_0(\varphi) + p_1(\varphi)\, r + p_2(\varphi)\, r^2$$
$$H(r, \varphi) = \varphi + q_0(\varphi) + q_1(\varphi)\, r + q_2(\varphi)\, r^2 \tag{6.2.1}$$

of map (1.1.4). It is assumed that $p_n \in P$ and $q_n \in P$, $0 \le n \le 2$ are given functions. Then G and H satisfy Assumption 2.2. Substituting (1) in the condition of invariance (1.1.7) yields

$$S(\varphi + q_0(\varphi) + q_1(\varphi) \cdot S(\varphi) + q_2(\varphi) \cdot (S(\varphi))^2) =$$
$$p_0(\varphi) + p_1(\varphi) \cdot S(\varphi) + p_2(\varphi) \cdot (S(\varphi))^2.$$

For $p_0(x) = 0$ it is seen that $S(\varphi) = 0$ is an invariant curve of (1). As a consequence we use $S_0(\varphi) = 0$ as a first initial approximation for Algorithm 1 (see Section 2.2) in most numerical experiments.

Let

$$\begin{aligned} p_0(\varphi) &= \alpha \cdot \sin\varphi,\ \alpha = 0.15, & q_0(\varphi) &= 0, \\ p_1(\varphi) &= 2, & q_1(\varphi) &= 1, \\ p_2(\varphi) &= -1, & q_2(\varphi) &= 0 \end{aligned} \tag{6.2.2}$$

in (1). By choosing the initial approximations

$$S_0(\varphi) = 0 \tag{6.2.3}$$

and

$$S_0(\varphi) = 1 \tag{6.2.4}$$

we observe that Algorithm 1 converges to two different invariant curves. The respective numerical experiments are described in Paragraphs A and E in this section. In Paragraph B and C the invariant curve computed with (3) is tested by choosing two additional initial approximations. Paragraph D is concerned with the realisation of the simplified Newton-Raphson method introduced in iteration (2.1.8).

The spline method formulated in Algorithm 2 (see Section 5.1) is used throughout in Paragraphs A to D. In step 1 of Algorithm 2 equidistant knots are applied. The number of knots is successively doubled. Working with a tolerance $\varepsilon = 10^{-5}$ the systems of linear equations are solved by Gauss elimination.

The following notation is adopted in the subsequent tables. Let in the n^{th} Newton-Raphson step v_n be one of the functions a_n, b_n, d_n, S_n considered in (2.2.2), (2.2.3), respectively. Furthermore let N_n^{max} be the maximum number of knots considered in step 1 of the inner loop in Algorithm 2. Then $\min|v_n|$, $\max|v_n|$, respectively denotes the minimum, maximum, respectively of the numerical values $|v_n(\varphi_j)|$, $\varphi_j \in r$, computed with the partition $r = \{x_0,\ldots,x_{N_n^{max}}\}$ given by (5.1.1).

A. The initial approximation (3)

We start the description of the numerical experiments by applying Algorithm 1 to the map (1) with (3). As in Example 1.3 (see Chapter 1 and Section 6.1) the 1^{st} Newton-Raphson step can be calculated explicitly. The evaluation of (2.2.2) with n = 0 and (1), (2) yields

$$h_0(\varphi) = \varphi,\ a_0(\varphi) = 2,\ b_0(\varphi) = \alpha \cdot \sin \varphi$$

and

$$d_0(\varphi) = \frac{b_0(\varphi)}{1 - a_0(\varphi)} = -\alpha \cdot \sin \varphi.$$

Furthermore we have with (2.2.3)

$$S_1(\varphi) = -\alpha \cdot \sin \varphi. \tag{6.2.5}$$

The computed spline function approximates (5) in the knots considered with full precision, thus the inner loop of Algorithm 2 stops with 8 knots (see Table 3). The coefficient functions of (2.2.2a) in the 2^{nd} Newton-Raphson step are also explicitly known. Again using (2.2.2) with n = 1 and (1), (2), (3), (5) we have

$$h_1(\varphi) = \varphi - \alpha \cdot \sin \varphi,$$

$$a_l(\varphi) = 2 + 2\alpha \cdot \sin\varphi + \alpha \cdot \cos(\varphi - \alpha \cdot \sin\varphi), \tag{6.2.6}$$

$$b_l(\varphi) = \alpha \cdot \sin(\varphi - \alpha \cdot \sin\varphi) - \alpha \cdot \sin\varphi - \alpha^2 \cdot \sin^2\varphi.$$

The characteristic exponents (4.2.4) are

$$\mu_0(S_1) = \frac{\log 2.15}{\log 0.85} = -4.71, \ \mu_\pi(S_1) = \frac{\log 1.85}{\log 1.15} = 4.40. \tag{6.2.7}$$

With (6) it is concluded that the linear functional equation

$$(d_l \circ h_l)(\varphi) = a_l(\varphi) \cdot d_l(\varphi) + b_l(\varphi) \tag{6.2.8}$$

satisfies the assumptions of Theorem 4.7. Thus (8) has a unique solution

$$d_l \in P^4.$$

We note the following:

1. Using induction on n and, assuming $a_n(0) \neq 1$ and $a_n(\pi) \neq 1$, it is shown with ease that

   $h_n(0) = 0$, $h_n(\pi) = \pi$, $S_n(0) = 0$, $S_n(\pi) = 0$, $n = 0, 1, 2, \ldots.,$

   i.e. h_n has a fixed point in $\varphi = 0$, $\varphi = \pi$, respectively and S_n has a zeros in $\varphi = 0$, $\varphi = \pi$, respectively. Furthermore by considering $n \to \infty$ we have

   $H(0, S(0)) = 0$, $H(\pi, S(\pi)) = \pi$ in (2.2.2a) and $S(0) = 0$, $S(\pi) = 0$

   (see Figures 4 and 5).

2. The positive characteristic exponent

   $$\mu_\pi(S_2) = \frac{\log a_2(\pi)}{\log \varphi_2'(\pi)}$$

   is approximated numerically. As $\varphi = \pi$ is a kot in our computation and $N_n^{max} = 128$ we consider

$$\mu_\pi^2(S_2) \approx \frac{\log a_2(\pi)}{\log \dfrac{h_2(\varphi_{63}) - h_2(\varphi_{65})}{\varphi_{63} - \varphi_{65}}}$$

with

$$\varphi_{63} = \frac{63\pi}{64} \text{ and } \varphi_{65} = \frac{65\pi}{64}$$

and find

$$\mu_\pi^2(S_2) \approx 3.48. \qquad (6.2.9)$$

Evaluating the condition of invariance (1.1.7) for the map (1) with (2) yields

$$\mu_\pi(S) = 3.53.$$

Table 3 summarises the results of the computation.

Newton-Raphson step n	$\max\lvert d_n \rvert$	N_n^{max}	μ_π
0	$1.5 \cdot 10^{-1}$	8	not defined
1	$3.5 \cdot 10^{-2}$	64	4.40 (see (7))
2	$2.0 \cdot 10^{-3}$	128	3.48 (see (9))

Table 3

B. A check of the implementation

We choose

$$S_0(\varphi) = 0.1 + 0.1 \cdot \sin\varphi \qquad (6.2.10)$$

as an additional initial approximation to (3) for Algorithm 1. We observe that the Newton-Raphson method converges towards the same invariant curve as the one found with (3). Table 4 summarises the results of the numerical experiment.

Newton-Raphson step n	$\max\|d_n\|$	N_n^{max}
0	0.47	64
1	0.15	128
2	$2.7 \cdot 10^{-2}$	128
3	$9.9 \cdot 10^{-4}$	128

Table 4

In Figure 6 the Newton-Raphson approximation S_n, $0 \le n \le 3$ are depicted. Sample of numerical values of the invariant curve are given in Table 5.

C. A nondifferentiable initial approximation

Let us consider

$$S_0(\varphi) = \begin{cases} \frac{0.2}{\pi}\varphi, & 0 \le \varphi \le \frac{\pi}{2} \\ -\frac{0.2}{3\pi}\varphi + \frac{0.4}{3}, & \frac{\pi}{2} \le \varphi \le 2\pi \end{cases}. \tag{6.2.11}$$

The Newton-Raphson approximations S_n, n = 0, 1, 2, 3, 4 are depicted in Figures 7, 8 and 9. We observe that the 1st Newton-Raphson approximations are rough. However, the Newton-Raphson method removes the nondifferentiability in $\varphi = \frac{\pi}{2}$ and converges towards the smooth invariant curve S.

Using the initial approximation (3), (10) and (11), Table 5 shows numerical approximations of the invariant curve in the knots

$$\varphi_k = k \cdot \frac{\pi}{4},\ 0 \le k \le 7.$$

k	(3)	(10)	(11)
0	0.00000'000	0.00000'059	0.00000'000
1	-0.08973'206	-0.08973'215	-0.08973' 178
2	-0.13054'862	-0.13054'865	-0.13054'861
3	-0.10323'483	-0.10323'482	-0.10323'482
4	0.00000'000	0.00000'000	0.00000'000
5	0.14062'005	0.14061'767	0.14061'775
6	0.17088'498	0.17088'487	0.17088'487
7	0.10396'929	0.10396'928	0.10396'929

Table 5

D. A simplified Newton-Raphson method

In the following the application of (2.1.8) to (1) with (3) is described. Assuming that an initial approximation $S_0^*(\varphi) = S_0(\varphi)$ is given, we solve

$$(d_n^* \circ h^*)(\varphi) - a^*(\varphi) \cdot d_n^*(\varphi) = b_n^*(\varphi),\ n = 0, 1, 2, \ldots. \qquad (6.2.12a)$$

for the unknown function d_n^*. The coefficient functions of (12) are given by

$$h^*(\varphi) = H(S_0(\varphi), \varphi),$$

$$a^*(\varphi) = G_r(S_0(\varphi), \varphi) - S_0'(H(S_0(\varphi), \varphi)) \cdot H_r(S_0(\varphi), \varphi),$$

$$b_n^*(\varphi) = - S_n^*(H(S_n^*(\varphi), \varphi)) + G(S_n^*(\varphi), \varphi).$$

The approximations of the invariant curve are then computed by

$$S_{n+1}^*(\varphi) = S_n^*(\varphi) + d_n^*(\varphi),\ n = 0, 1, 2, \ldots.. \qquad (6.2.12b)$$

Contrary to the Newton-Raphson method formulated in Algorithm 1, the coefficient functions h^* and a^* in (12a) are not adjusted to the n^{th} Newton-Raphson approximation. Using the initial approximation S_0^* they are kept constant in the simplified Newton-Raphson method (2.1.8), (12), respectively. Based on (5) we choose

$$S_0^*(\varphi) = - \alpha \cdot \sin \varphi$$

and for the numerical realisation the tolerance is set to $\varepsilon = 10^{-5}$. We observe convergence of Algorithm 2 with systems of linear equations of order 128. In Table 6 the convergence of the corrections d_n^* is shown.

Newton-Raphson step n	$\max\left|\tilde{d}_n\right|$	N_n^{max}
0	$3.5 \cdot 10^{-2}$	128
1	$1.9 \cdot 10^{-3}$	128
2	$2.1 \cdot 10^{-4}$	128
3	$1.9 \cdot 10^{-5}$	128
4	$1.7 \cdot 10^{-5}$	128

Table 6

As expected, the convergence of the simplified Newton-Raphson method (12) is not quadratic. However, the comparison of the computed values of the invariant curve with the figures in Table 5 shows that the requested 5 digit accuracy is achieved.

E. The initial approximation (4)

We choose (4) as initial approximation in Algorithm 1 and report of the numerical results computed by using Newton-Raphson methods formulated in Algorithm 1. We observe that Algorithm 1 converges towards another invariant curve than the one found with (3). We use the Fourier method and find convergence in 3 Newton-Raphson steps. The invariant curve is represented with Fourier series of length 32. Figures 10 and 11 show the invariant curve S and the function H given by (1) and evaluated with the numerical values of S.

In Table 7 we give the minima and the maxima of the coefficient functions $a_n(\varphi)$, n = 0, 1, 2, 3 defined by (2.2.2c) and computed by using the initial approximations (3) and (4). In Figures 12 and 13 the function a on the invariant curve defined by

$$a(S)(\varphi) = G_r(S(\varphi), \varphi) - S'(H(S(\varphi), \varphi)) \cdot H_r(S(\varphi), \varphi)$$

(see (2.2.2c)) and evaluated with the numerical values of S is depicted.

	$S_0(\varphi) = 0$		$S_0(\varphi) = 1$	
n	$\min\lvert a_n \rvert$	$\max\lvert a_n \rvert$	$\min\lvert a_n \rvert$	$\max\lvert a_n \rvert$
0	2.00	2.00	0.00	0.00
1	1.68	2.35	$3.4 \cdot 10^{-4}$	0.21
2	1.63	2.28	$1.4 \cdot 10^{-4}$	0.21
3	1.63	2.28	$1.4 \cdot 10^{-4}$	0.21

Table 7

As the Newton-Raphson method in one dimension is converging for attractive and repelling fixed points, the Newton-Raphson method considered in this thesis is converging for $\lvert a_n(\varphi) \rvert < 1$ and $\lvert a_n(\varphi) \rvert > 1$, $\forall \varphi \in \mathbf{R}^1$, n = 0, 1, 2,….. The fixed point iteration (1.4.7), however, can only be used for computing the invariant curve found with the initial approximation (3).

In Table 8 the minima and the maxima of the computed values of the Newton-Raphson approximations S_n, n = 0, 1, 2,…. are given. With (1), (3) and (2.2.2b) we have

$$h_n(\varphi) = H(S_n(\varphi), \varphi) = \varphi + S_n(\varphi),\ n = 0, 1, 2, \ldots.$$

Thus, from the values in Table 8 it is seen that h_n, $0 \le n \le 3$ has no fixed points. As already reported in Section 6.1 the Fourier method approximates invariant curves that do not contain fixed points of the map (1.1.4) with high accuracy. However, the fast convergence of the Fourier series indicates that either the rotation number of the invariant curve under the given map is either irrational (see Theorem 4.10) or has a high degree of differentiability (see Theorem 4.7 and also Chapter 8).

Newton-Raphson step n	$\min\lvert S_n \rvert$	$\max\lvert S_n \rvert$
0	1.00	1.00
1	0.85	1.15
2	0.84	1.14
3	0.84	1.14

Table 8

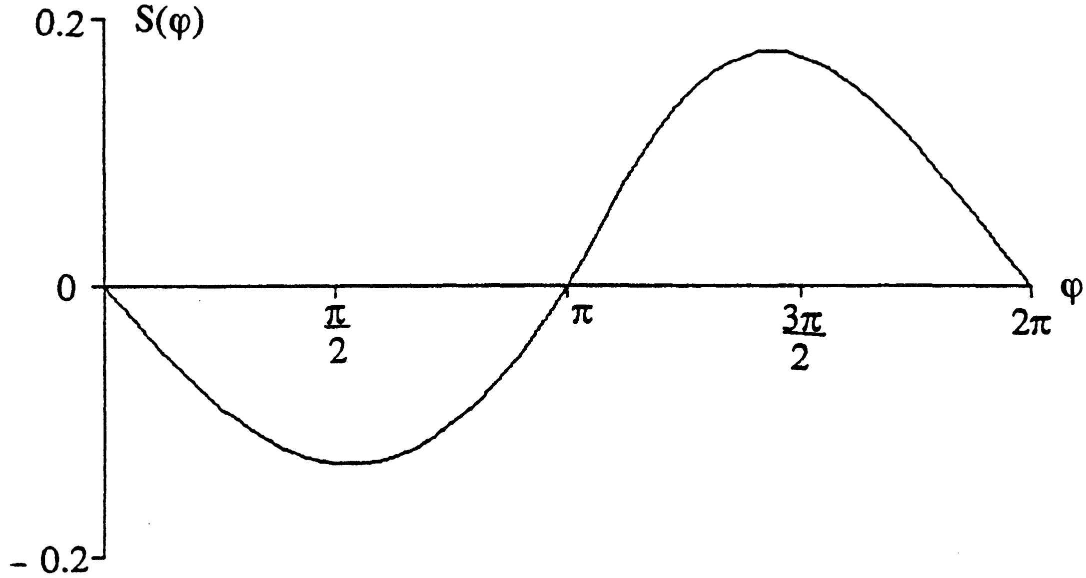

Figure 4

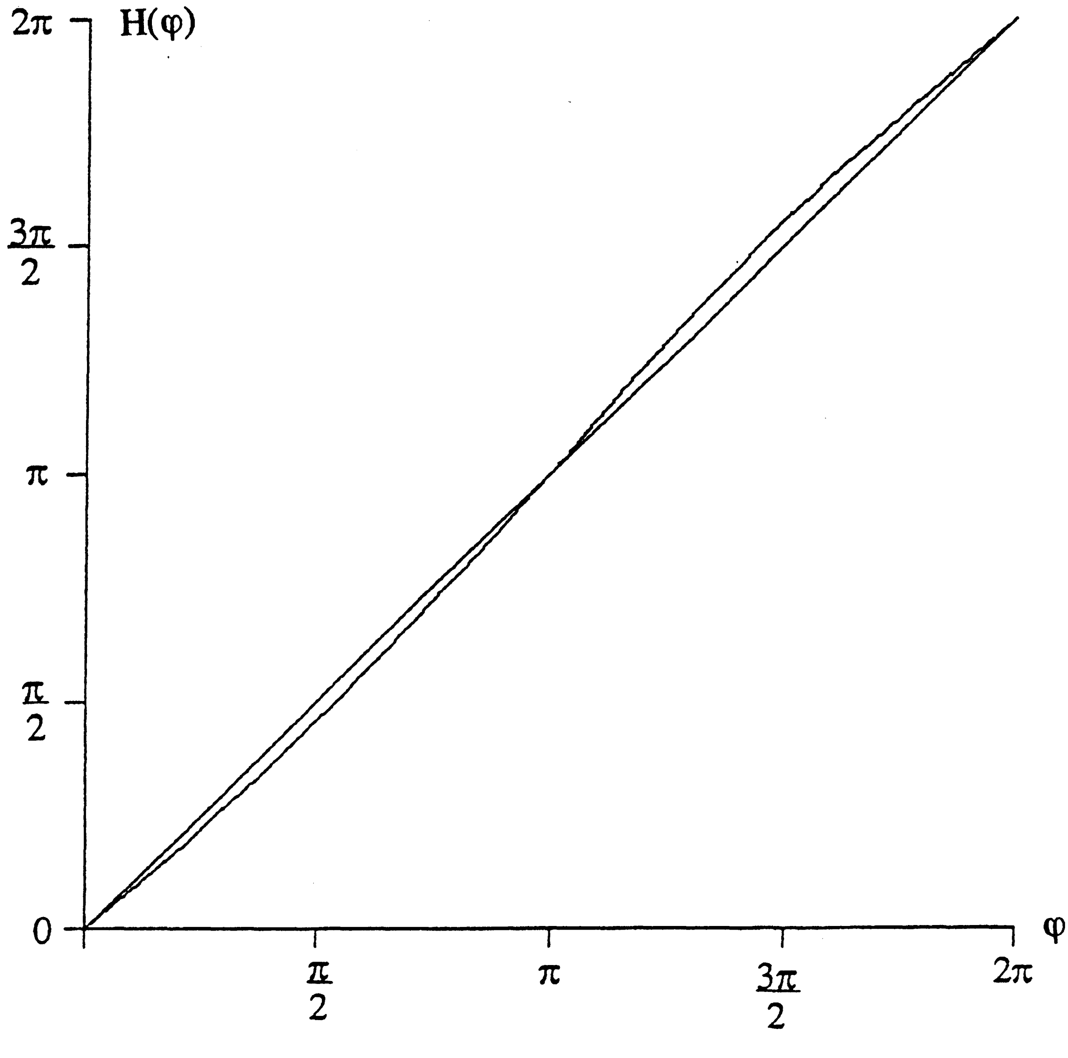

Figure 5

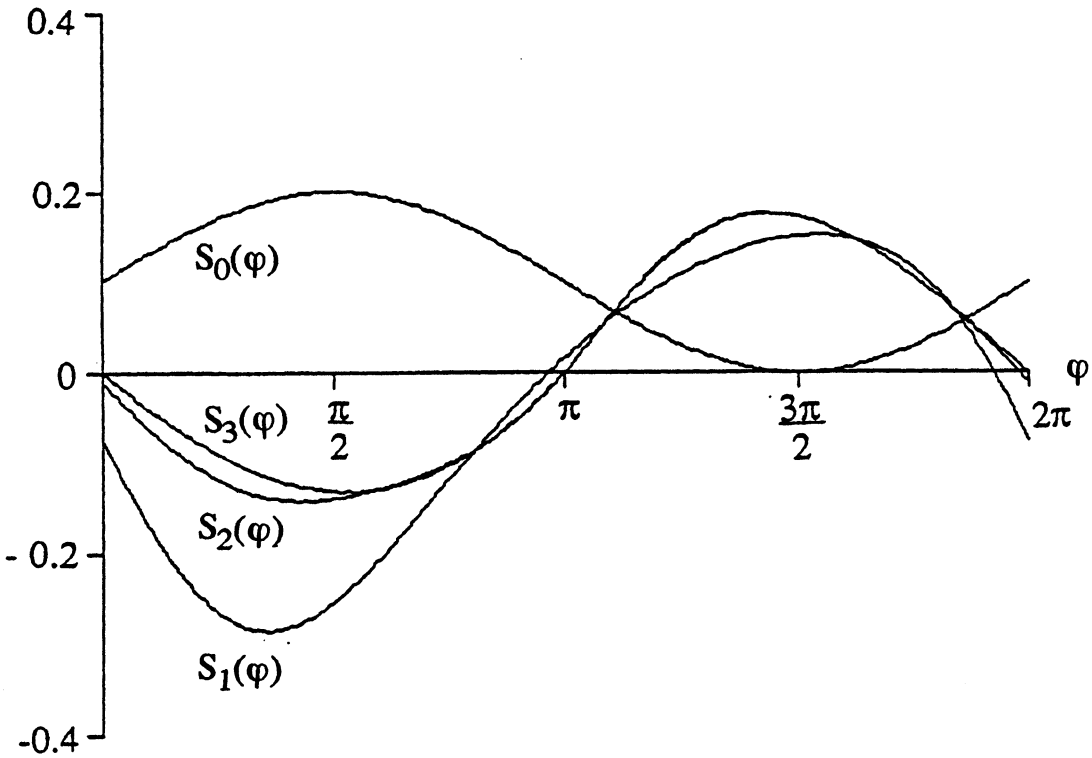

Figure 6

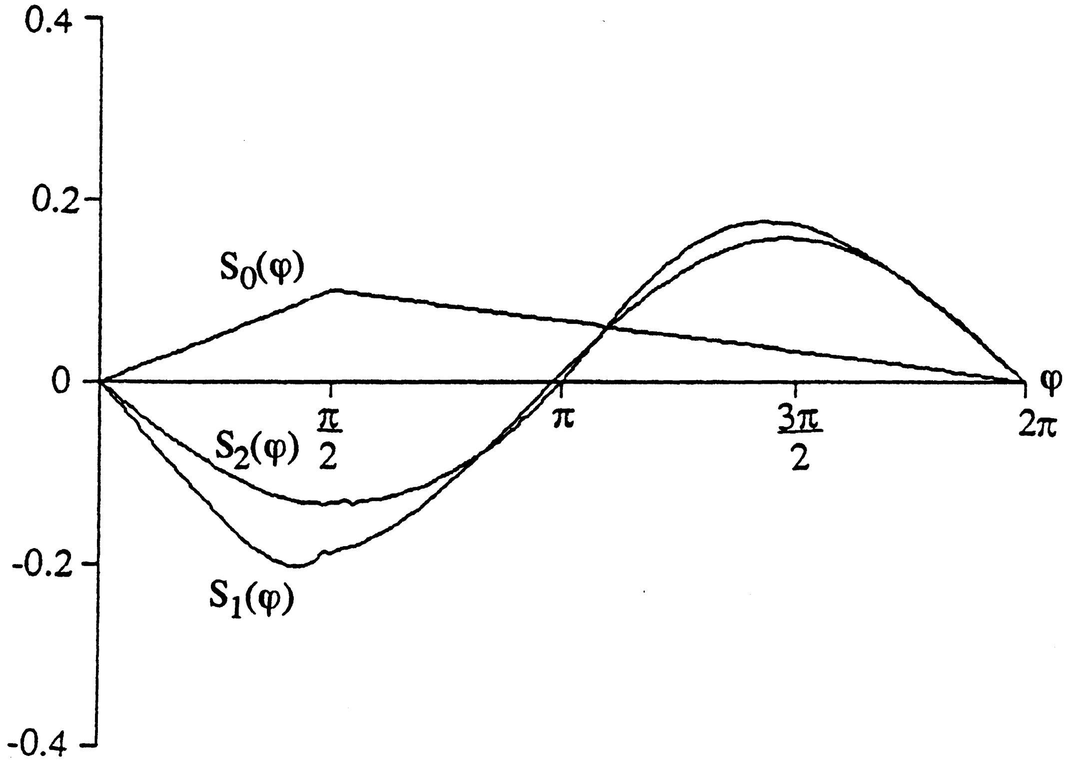

Figure 7

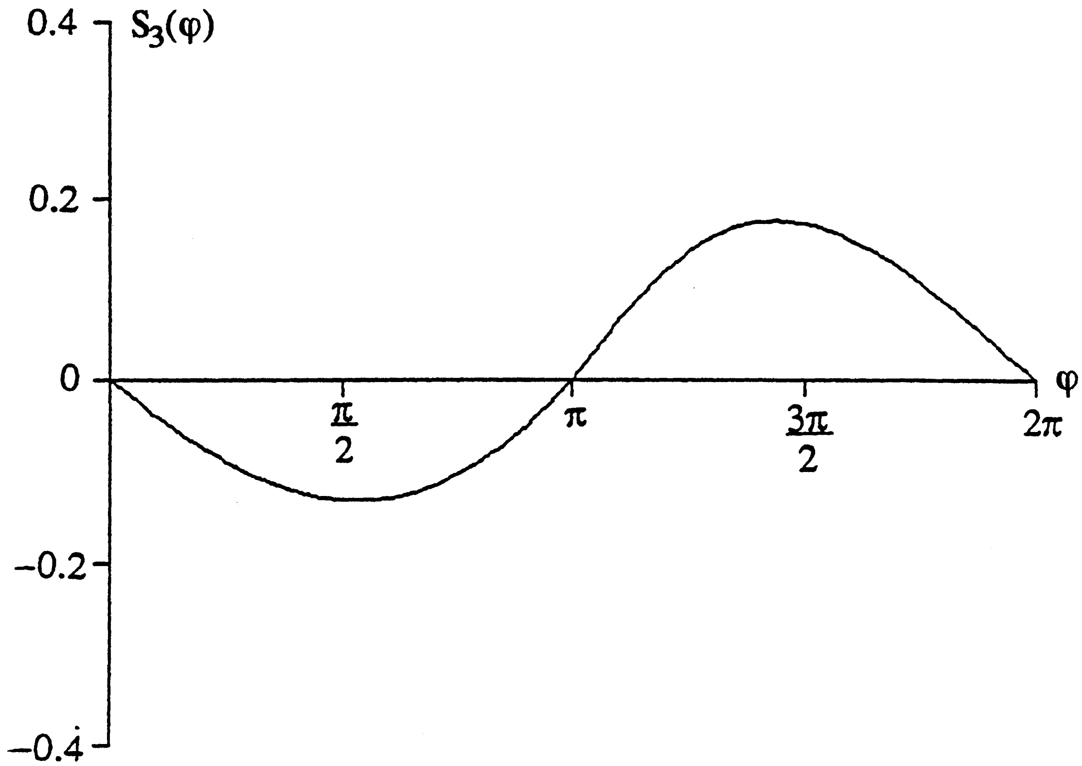

Figure 8

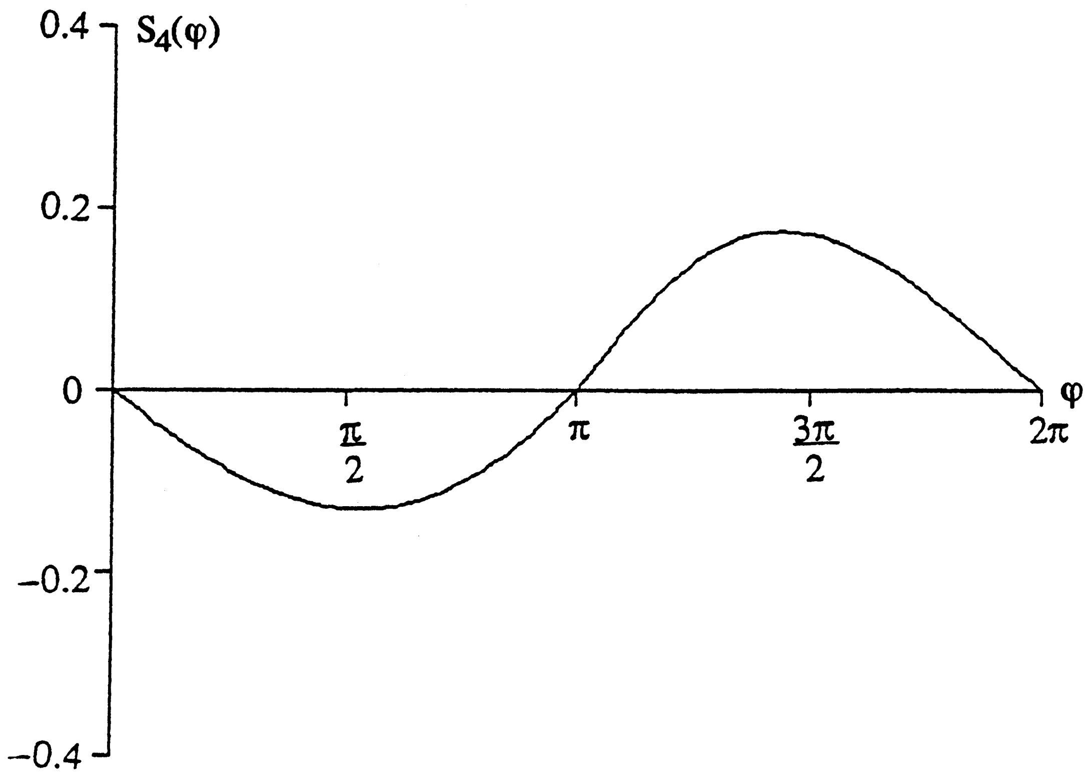

Figure 9

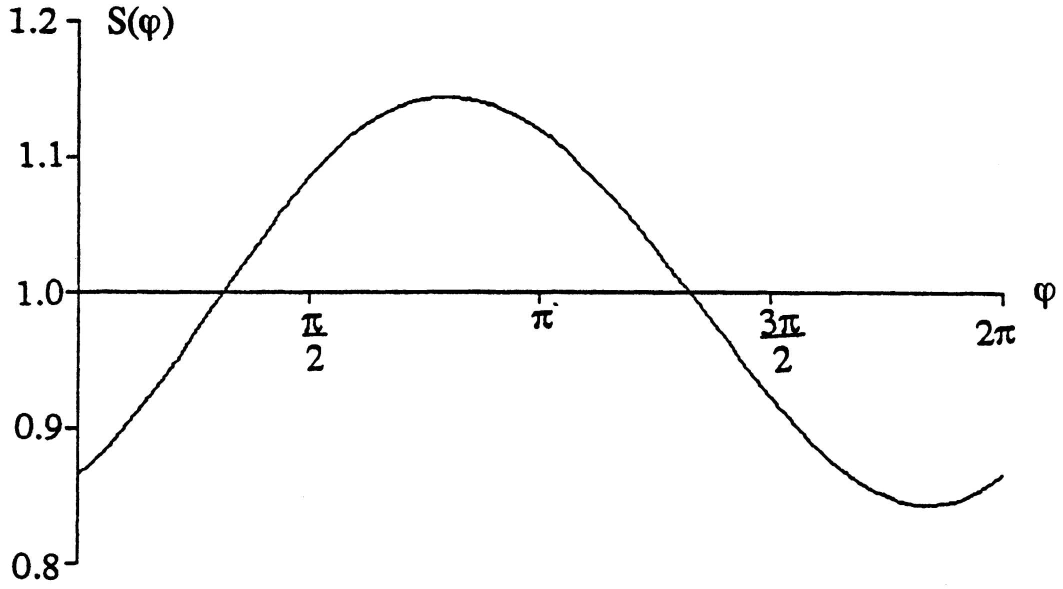

Figure 10

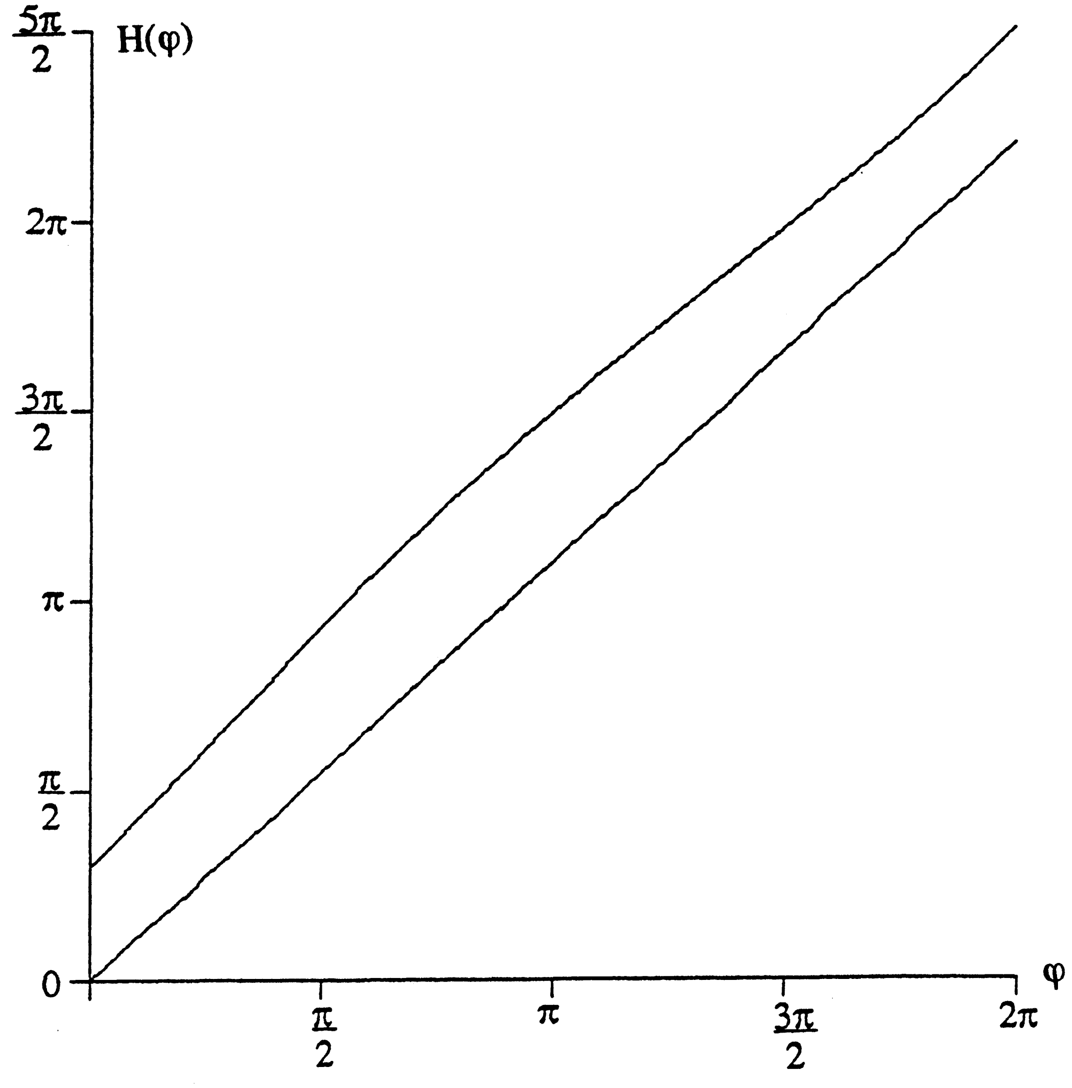

Figure 11

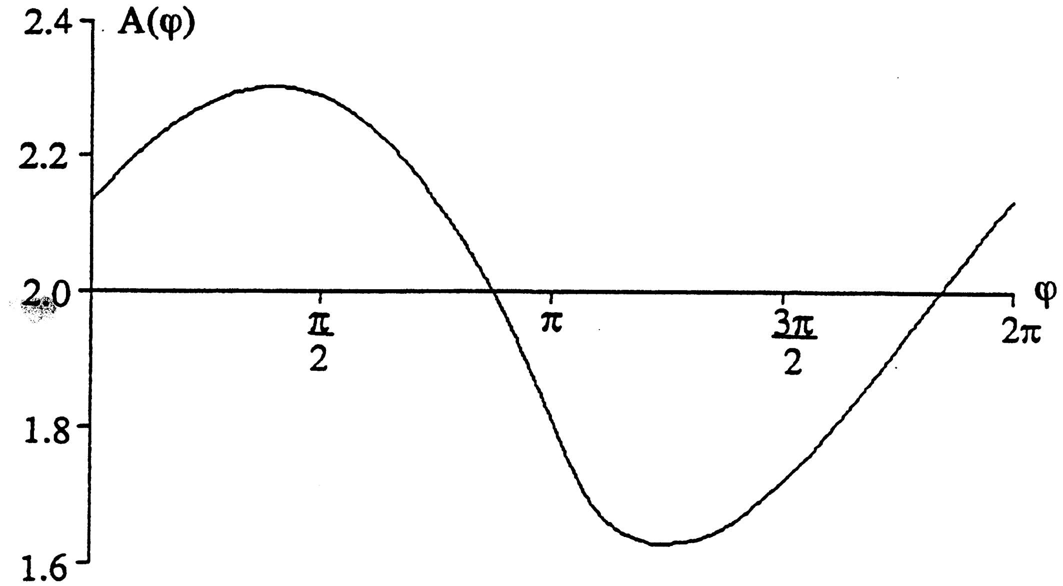

Figure 12

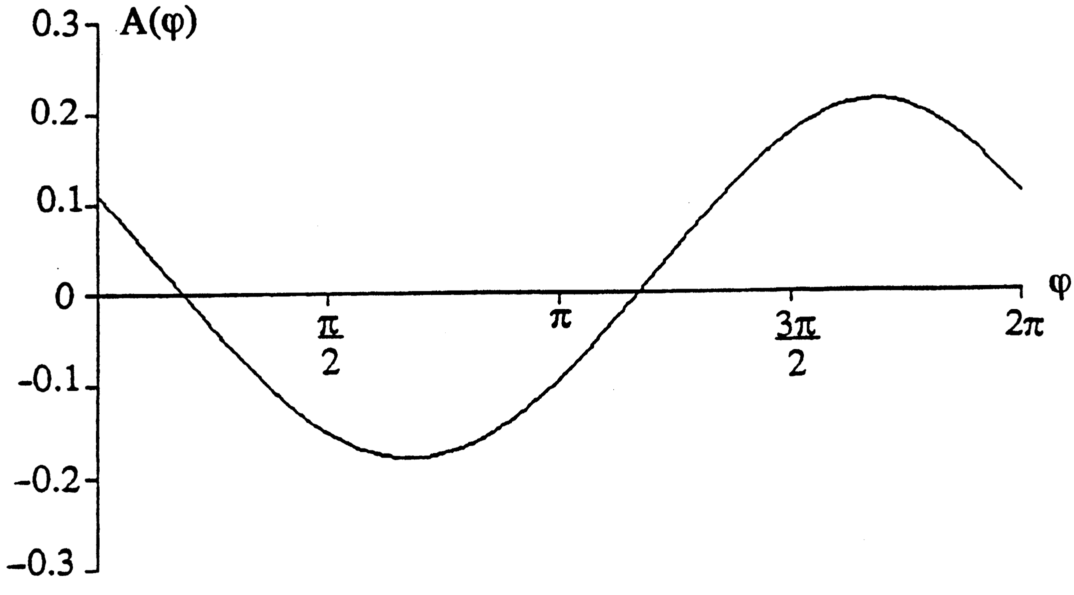

Figure 13

6.3. Splines versus Fourier series

We consider

$$
\begin{aligned}
p_0(\varphi) &= \frac{1}{b - c\sin\varphi - d\cos\varphi}, & q_0(\varphi) &= a\sin\varphi, \\
p_1(\varphi) &= e - d\sin\varphi, & q_1(\varphi) &= -\frac{1}{b - c\sin\varphi - d\cos\varphi}, \quad (6.3.1) \\
p_2(\varphi) &= -\frac{1}{b - c\sin\varphi - d\cos\varphi}, & q_2(\varphi) &= \frac{1}{b - c\sin\varphi - d\cos\varphi}
\end{aligned}
$$

in (6.2.1) where the constants are chosen as $a = -0.84$, $b = 3$, $c = 0.1$, $d = 0.2$ and $e = 2.3$. As assumed in (6.2.1) we have $p_n \in P$ and $q_n \in P$, $0 \le n \le 2$ in (1).

In this section we describe 3 different numerical experiments with (1). In Paragraph A the spline method (Algorithm 2) with equidistant knots (5.1.1) is used. Although the Newton-Raphson method converges with 8 digit precision, it is shown numerically that the accuracy of the computed values of the invariant curve S in the neighbourhood of the fixed point with positive characteristic exponent is considerably less. We improve the accuracy by using nonequidistant knots. In C the Fourier method (see Section 5.2) is also applied to the map (1). We compare Gauss elimination to the Algorithm 4 and 5 introduced in Section 5.3.

A. The spline method with equidistant knots

Starting from the initial approximation

$$S_0(\varphi) = 0 \qquad (6.3.2)$$

we compute Newton-Raphson approximations by choosing $N_n^{max} = 128$ in Algorithm 2. In Table 9 the values

$$\max\left|\tilde{S}_{n+1} - \hat{S}_{n+1}\right| \text{ and } \max\left|d_n\right|$$

are shown.

Newton-Raphson step n	$\max\left\|\tilde{S}_{n+1} - \hat{S}_{n+1}\right\|$	$\max\left\|d_n\right\|$
0	$1.6 \cdot 10^{-5}$	$3.1 \cdot 10^{-1}$
1	$3.3 \cdot 10^{-4}$	$3.0 \cdot 10^{-2}$
2	$1.5 \cdot 10^{-5}$	$2.7 \cdot 10^{-4}$
3	$3.1 \cdot 10^{-8}$	$8.1 \cdot 10^{-8}$
4	$6.0 \cdot 10^{-8}$	$1.5 \cdot 10^{-7}$

Table 9

The invariant curve S is depicted in Figure 15 and Figure 16 shows the map H evaluated with S. By choosing a second initial approximation

$$S_0(\varphi) = 0.1 + 0.1 \cdot \sin\varphi \tag{6.3.3}$$

the computed values of S are checked. Using the initial conditions (2) and (3), Table 10 shows the fixed points s and t of h_n, $1 \le n \le 4$ defined by (2.2.2b).

	$S_0(\varphi) = 0$		$S_0(\varphi) = 0.1 + 0.1 \cdot \sin\varphi$	
	s	t	s	t
0	0.00000'00000	3.14159'26535	6.24745'50765	3.17509'22598
1	6.20387'19919	3.21175'76628	6.20257'58137	3.21418'54214
2	6.20025'92229	3.21543'44310	6.20025'03719	3.21547'18050
3	6.20024'00137	3.21547'27023	6.20023'00133	3.21547'26994
4	6.20024'00152	3.21547'26994	6.20024'00154	3.21547'26994

Table 10

We note that using (2) we have for n = 0 with (1) and (6.2.1)

$$h_0(\varphi) = \varphi - 0.84 \cdot \sin\varphi$$

and

$$\mu_s(S_0) = \frac{\log 2.30}{\log 0.16} = -0.45,\ \mu_t(S_0) = \frac{\log 2.30}{\log 1.83} = 1.37,$$

i.e. the fixed point of h_0 are in 0 and π and the characteristic exponents are explicitly known. As we introduce equidistant knots on $[0, 2\pi]$ in Algorithm 2, the fixed points of h_0 are elements of the partition (5.1.1). The rest of the fixed point values shown in Table 10 do not belong to partitions considered in Algorithm 2. As described in Section 5.1, we represent the Newton-Raphson approximation by a 2π-periodic spline function and, using a one dimensional Newton-Raphson method, we compute the fixed point of $h_n(\varphi)$, $0 \le n \le 4$, defined by (2.2.2b). Thus we approximate

$$\mu_s(S_n) \approx \frac{\log a_n(s)}{\log h'_n(s)} \text{ and } \mu_t(S_n) \approx \frac{\log a_n(t)}{\log h'_n(t)},\ 1 \le n \le 4.$$

Nonequidistant knots (see Paragraph B) and the initial approximation (1) yield the numerical values

Newton-Raphson step n	μ_s^n	μ_t^n
1	-0.44	1.23
2	-0.40	1.22
3	-0.40	1.19
4	-0.40	1.19

Table 11

Based on (1.1.7), the condition for a fixed point of H that is an element of S is given by

$$\begin{aligned} &H(S(\varphi), \varphi) - \varphi = 0 \\ &G(S(\varphi), \varphi) - S(H(S(\varphi), \varphi)) = 0. \end{aligned} \tag{6.3.4}$$

We compute the fixed points s and t on the invariant curve with machine precision and find

$$s = 6.20024'00137,\ t = 3.21542'89014,$$

and the interpolation of S in s and t yields

$$Y(s) = 0.28187'99493,\ Y(t) = 0.26028'17694.$$

In addition, by differentiation of (4) we find for the characteristic exponents

$$\mu_s = -0.40,\ \mu_t = 1.19.$$

The following is concluded:

1. Comparing the values in (4) and Table 10 it is seen that the Newton-Raphson method approximates the fixed points s and t of $H(S(\varphi), \varphi)$ with different accuracies.

2. Following Theorem 4.7 it is indicated that the Newton-Raphson approximations and the invariant curve are in P^l.

B. The spline method with nonequidistant knots

In Paragraph A it is shown that, using equidistant knots, Algorithm 1 converges towards the "wrong" fixed point t. As illustrated in Figure 16, the spline function approximates the invariant curve with lower accuracy than the precision specified by the Newton-Raphson method in Algorithm 1.

We introduce more knots in the neighbourhood of t by considering the function

$$g(\varphi) = \varphi - \sin(\varphi - 3.2154) \tag{6.3.5}$$

and the partitions

$$t = \{g(\varphi_0),\ldots,g(\varphi_{63})\} \text{ and } r = \{g(\varphi_0),\ldots,g(\varphi_{127})\}$$

with

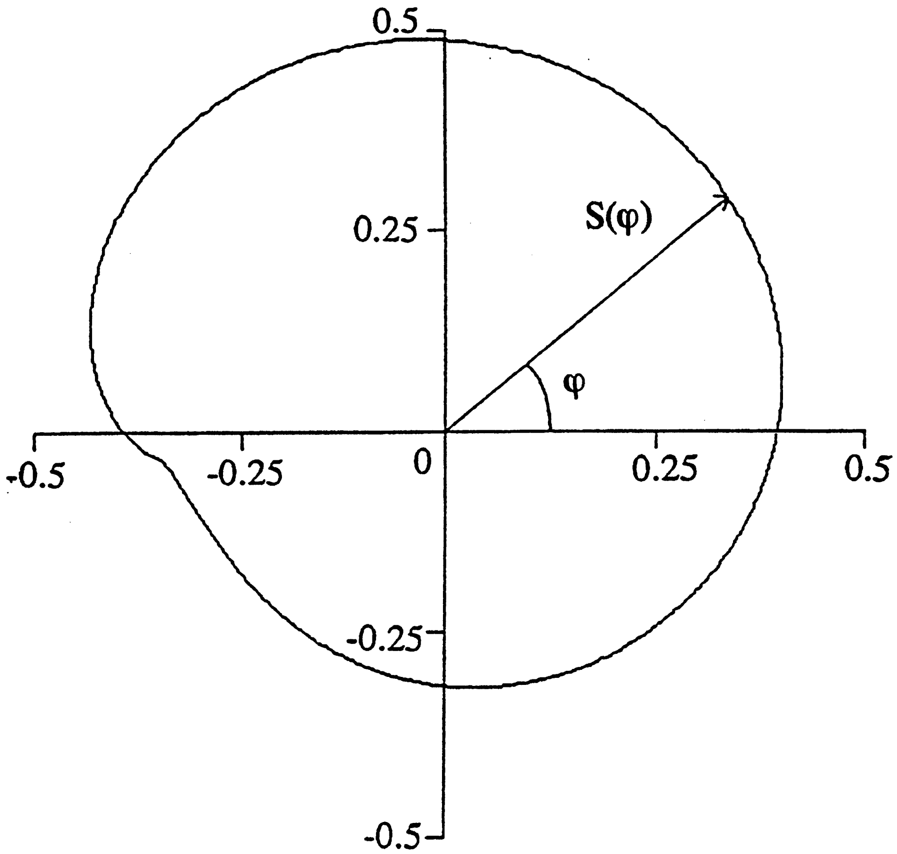

Figure 14

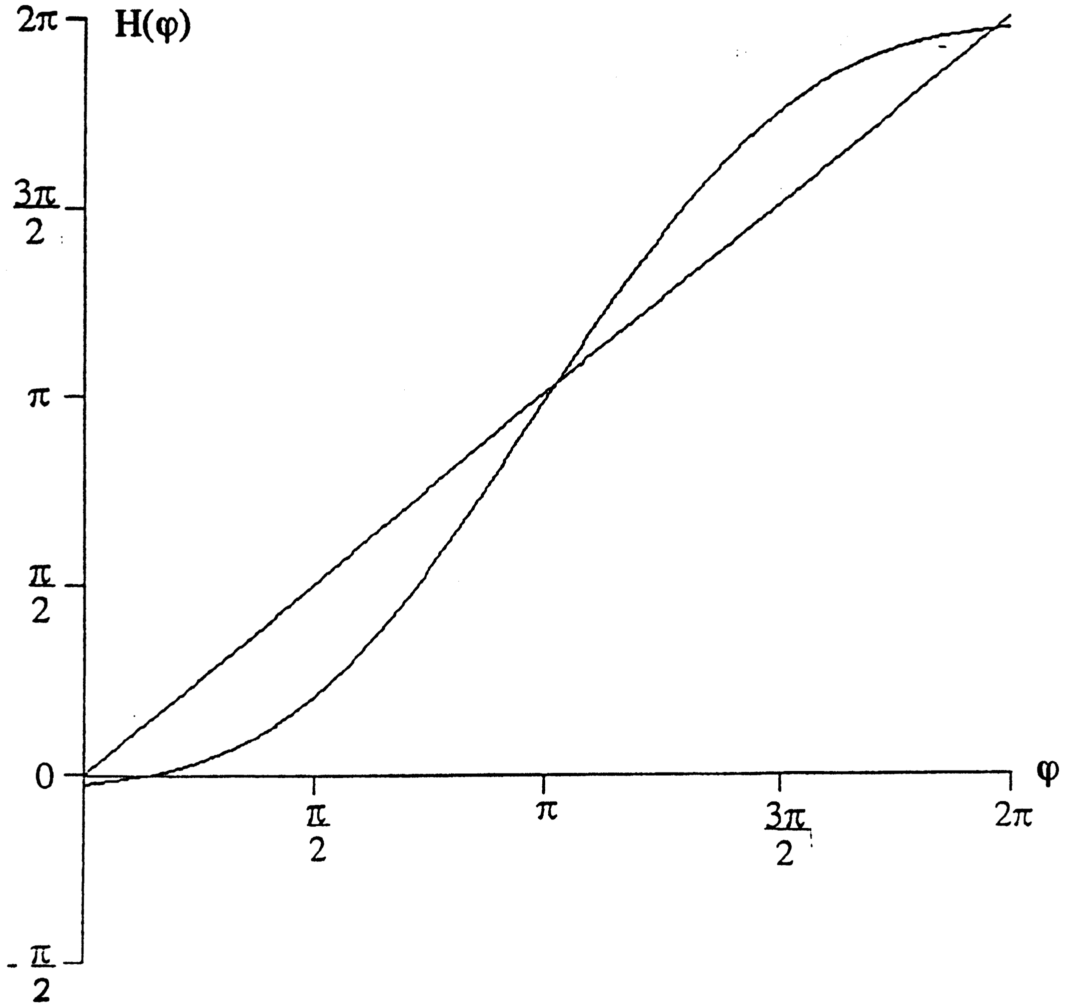

Figure 15

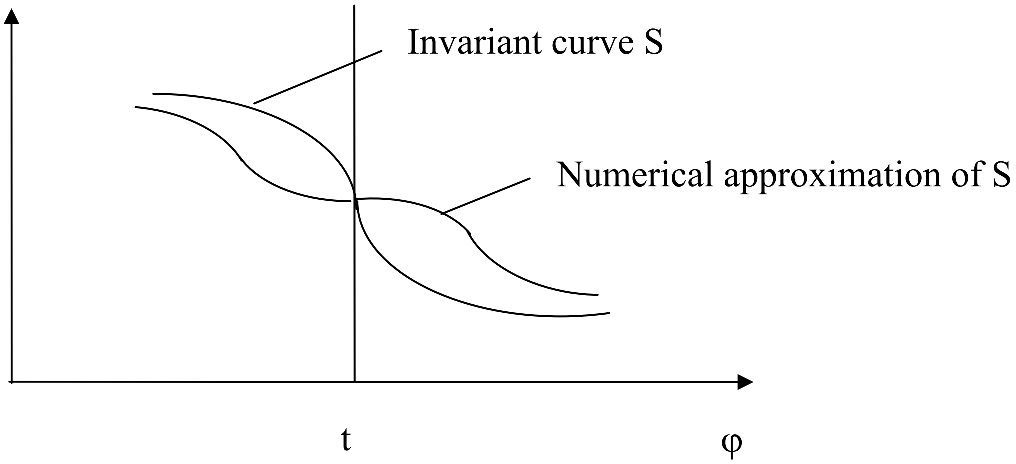

Figure 16

$$\varphi_j = \frac{2\pi}{N} \cdot j,\ j = 0,\ldots,N-1,\ N = 64,\ N = 128 \qquad (6.3.6)$$

in Algorithm 2. The function g defined by (5) concentrates the equidistant knots (6) in $\varphi = 3.2154$. As in Table 9 we note in Table 12 the values

$$\max\left|\tilde{S}_{n+1} - \hat{S}_{n+1}\right| \text{ and } \max\left|d_{n+1}\right|$$

computed in Algorithm 2.

initial approximation	$S_0(\varphi) = 0$		$S_0(\varphi) = 0.1 + 0.1 \cdot \sin\varphi$	
n	$\max\left\|\tilde{S}_{n+1} - \hat{S}_{n+1}\right\|$	$\max\left\|d_n\right\|$	$\max\left\|\tilde{S}_{n+1} - \hat{S}_{n+1}\right\|$	$\max\left\|d_n\right\|$
0	$1.5 \cdot 10^{-5}$	$3.0 \cdot 10^{-1}$	$6.6 \cdot 10^{-6}$	$1.8 \cdot 10^{-1}$
1	$3.7 \cdot 10^{-5}$	$3.0 \cdot 10^{-2}$	$6.5 \cdot 10^{-6}$	$1.3 \cdot 10^{-2}$
2	$1.8 \cdot 10^{-5}$	$4.2 \cdot 10^{-4}$	$4.9 \cdot 10^{-6}$	$9.5 \cdot 10^{-5}$
3	$1.8 \cdot 10^{-6}$	$2.0 \cdot 10^{-6}$	$1.5 \cdot 10^{-6}$	$1.7 \cdot 10^{-7}$
4	$4.2 \cdot 10^{-9}$	$4.7 \cdot 10^{-9}$	$7.0 \cdot 10^{-10}$	$6.6 \cdot 10^{-11}$

Table 12

In order to test the numerical results in Paragraph A and B we choose a partition (5.1.1) not used in Algorithm 2, apply the map (1) and check the condition of invariance (1.1.7). Thus we consider

$$\varphi_j = \frac{\pi}{50} \cdot j,\ j = 0,\ldots,99 \tag{6.3.7}$$

with

$$b(\varphi_j) = S(H(S(\varphi_j), \varphi_j)) - G(S(\varphi_j), \varphi_j),\ j = 0,\ldots, 99.$$

Then, the number digits with which the invariant curve S satisfies the condition of invariance (1.1.7) in the knots φ_j, $j = 0,\ldots, 99$ is given by

$$-\log\left|b(\varphi_j)\right|.$$

Table 13 summarises the results of the computation.

j in (7)	equidistant knots (6) with (1)	nonequidistant knots with (1) by using (5)
1-45	8-10	8-9
46	7	10
47	6	8
48	7	9
49	7	8
50	6	8
51	5	7
52	4	9
53	4	7
54	5	7
55	4	9
56	4	7
57	5	7
58	6	8
59	7	7
60	7	7
61	7	8
62	7	8
63-100	8-10	8-10

Table 13

Table 13 shows the expected result: By concentrating the knots in the neighbourhood of the fixed point t the accuracy of the computed values of the invariant curve is considerably improved. The design of an algorithm that automatically concentrates the number of knots close to fixed points with small positive characteristic exponents is left to future research.

C. Comparison of Algorithm 4 and 5 (see Section 5.3) with Gauss elimination

The Fourier method (see Section 5.2) can also be applied to map (6.2.1) with the coefficient functions (1). Following (5.2.13) and Algorithm 3, let us consider systems of linear equations

$$A\,\delta = b,\ A \in \mathbf{R}^{N x N},\ b \in \mathbf{R}^{N}$$

for the unknown vector of Fourier coefficients $\delta \in \mathbf{R}^N$ and varying dimension N. Working with a tolerance $\varepsilon = 10^{-6}$, Algorithm 3 converge with systems of linear equations of order N = 1024. We observe that the computer time for generating the Matrix A in (5.2.13) and solving a system of linear equation with dimension N changes little between the different Newton-Raphson steps. Algorithms 4 and 5 are used for solving these large systems.

In the following we compare the iterative solver introduced in Section 5.3 with Gauss elimination. The implementation of Algorithm 4 and 5 is discussed and we show used computer time of the different algorithms.

1. The choice of the initial approximation in Algorithms 4 and 5

In Algorithm 3 systems of linear equations with increasing dimension have to be solved. The computed solutions are used as an initial approximation for the system with next larger dimension. By adding zeros the dimension of the computed solution is increased (see Figure 3).

2. The computation of the operation matrix times vector in Algorithms 4 and 5

The computation of A in (5.2.13) requires the evaluation of N basis vectors defined by (5.2.7) followed by N fast Fourier transforms. The values in Table 14 show that for large systems the computer time for computing the matrix A is substantial. In addition, if the implementation is required to avoid the storage of A, the values in Table 14 have to be multiplied by the number of iterations in Algorithms 4 and 5.

We propose a different method. Following (2.3.3) we consider the linear operator

$$L_{S_n}(d_n) = d_n \circ h_n - a_n \cdot d_n \tag{6.3.8}$$

where n denotes the n^{th} Newton-Raphson step in Algorithm 1. As the matrix A approximates the operator L by Fourier coefficients, we transform the vector of Fourier coefficients to function values, apply L given by (8) and transform back to Fourier coefficients. Based on the 2^{nd} Newton-Raphson step, Table 14 shows that this algorithm is more efficient for computing the operation matrix times vector in Algorithms 4 and 5.

Dimension N of A	Computer time for computing A [sec]	Computer time for computing Ax_k [sec]
4	$5 \cdot 10^{-3}$	$1 \cdot 10^{-2}$
8	$1 \cdot 10^{-2}$	$2 \cdot 10^{-2}$
16	$3 \cdot 10^{-2}$	$4 \cdot 10^{-2}$
32	$9 \cdot 10^{-2}$	$1 \cdot 10^{-2}$
64	$3 \cdot 10^{-1}$	$2 \cdot 10^{-2}$
128	1	$3 \cdot 10^{-1}$
256	5	$9 \cdot 10^{-1}$
512	18	2
1024	73	7

Table 14

3. Computing the Fourier coefficient of the Newton-Raphson correction in Algorithm 1

Using Algorithm 3 with the initial condition (2) we compare in Table 15 the iterative solvers with Gauss elimination. A_0 in Algorithm 4 is chosen with 2 diagonal below and beyond the main diagonal (see Figure 2). In Table 15 we again consider the 2^{nd} Newton-Raphson step. The computer time in the 2^{nd} column of Table 15 does not include the generation of A_0 in Algorithm 4. However, the implementation of Algorithm 4 requires the computation of the matrix A because we need A_0. Thus, the time for computing the solution of (1) by the iteration (5.3.2) consists of the values in the 2nd column of Table 14 and the 2nd column of Table 15.

N	Algorithm 4 [sec]	number of iterations	Algorithm 5 [sec]	number of iterations	Gauss elimination
4	no iteration	-	$1 \cdot 10^{-1}$	5	$<10^{-3}$
8	$2 \cdot 10^{-1}$	9	$4 \cdot 10^{-1}$	9	$1 \cdot 10^{-3}$
16	$4 \cdot 10^{-1}$	10	1	11	$5 \cdot 10^{-3}$
32	$9 \cdot 10^{-1}$	11	2	10	$2 \cdot 10^{-3}$
64	2	11	3	9	$1 \cdot 10^{-1}$
128	4	11	5	7	$7 \cdot 10^{-1}$
256	9	10	6	4	5
512	21	9	17	4	36
1024	70	10	35	3	289

Table 15

We summarize:

1. As we are working with low tolerance ($\varepsilon = 10^{-6}$) in the numerical experiments described in this paragraph it is seen in Table 15 that iterative solvers are more efficient than Gauss elimination.

2. We see that Algorithm 5 is considerably faster than Algorithm 4.

7. The Van-der-Pol equation

7.1. Basic properties

The system of ordinary differential equation

$$\dot{r} = \alpha \sin\varphi \cdot \{ -(1 - r^2 \cos^2\varphi)\, r \sin\varphi + c \cos \nu t\}$$
$$\dot{\varphi} = 1 + \frac{\alpha}{r} \cos\varphi \cdot \{ -(1 - r^2 \cos^2\varphi)\, r \sin\varphi + c \cos \nu t\},\ c \neq 0 \tag{7.1.1}$$

for the unknown function r(t) and φ(t) is called the *forced Van-der-Pol oscillator.* The coefficients $\alpha \in \mathbf{R}^1$, $c \in \mathbf{R}^1$ and $\nu \in \mathbf{R}^1$ in (1) are assumed to be given. Originally, Van-der-Pol derived (1) from an application in physics [26]. As (1) is discussed in the literature in great detail, see e.g. [12], [18] and [27], we limit our considerations to the description of the relation between (1) and the map (1.1.4).

We denote with $\varphi(t, r_0, \varphi_0)$, $r(t, r_0, \varphi_0)$ the solution of (1) with initial condition $r_0 \in \mathbf{R}^1$, $\varphi_0 \in \mathbf{R}^1$, i.e.

$$r(0, r_0, \varphi_0) = r_0$$

$$\varphi(0, r_0, \varphi_0) = \varphi_0.$$

With $r_0 \in \mathbf{R}^1$, $\varphi_0 \in \mathbf{R}^1$ as originals and $r \in \mathbf{R}^1$, $\varphi \in \mathbf{R}^1$ as images the *poincaré map* for $T \in \mathbf{R}^1$ and $T \neq 0$ is defined by

$$\tilde{r} = r(T, r_0, \varphi_0)$$
$$\tilde{\varphi} = \varphi (T, r_0, \varphi_0). \tag{7.1.2}$$

Using (1.1.4) and assuming that T is constant, a map $\mathbf{R}^2 \to \mathbf{R}^2$ is then given by

$$G(r_0, \varphi_0) = r(T, r_0, \varphi_0)$$
$$H(r_0, \varphi_0) = \varphi(T, r_0, \varphi_0). \tag{7.1.3}$$

In Kirchgraber [12, page 228] it is shown that the map (3) satisfies the periodicity condition (1.1.6).

In the following we discuss the dynamics of the solutions (2). Let ν_0 be the plane that contains the initial conditions r_0, φ_0 and let the t-axis be perpendicular to ν_0. From the averaging method (see e.g. Kirchgraber [12, 1978]) follows:

a. For $\alpha \to 0$, the system (1) has in the neighbourhood of the origin an asymptotically stable equilibrium solution, i.e. the map (3) has a fixed point and there exists an area such that every solution starting in that area converges towards this fixed point.

b. If $c \neq 0$ there exists in ν_0 a simple closed invariant curve S with

$$\left| S(r_0) - 2 \right| = O(\alpha),$$

more precisely by integrating (1) with

$$T = k\,\frac{2\pi}{\nu},\ k \in \mathbf{Z} \setminus \{O\} \tag{7.1.4}$$

there exists a curve S that is reproduced on planes with distances

$$\nu_k = k\,\frac{2\pi}{\nu},\ k \in \mathbf{Z} \setminus \{O\}$$

from ν_0. Thus, S is an invariant curve in ν_0 if we choose $k \in \mathbf{Z}$ with (4) and consider the perpendicular projection to ν_0. The planes ν_k contain copies of S. It is seen that in the algorithm for computing S we have to choose

$$T = k\,\frac{2\pi}{\nu},\ k \in \mathbf{Z} \setminus \{O\}.$$

c. S delimits the area of attraction of the equilibrium solution described in the previous remark a. S separates the bounded from the unbounded solutions of (1).

7.2. Numerical examples

For the computation of the invariant curve described in Section 7.1 by Algorithm 1 (see Section 2.2), we apply Algorithm 2 (see Section 5.1) for solving the linear sub problem (2.2.2a). We note that the computation of G_r and H_r in (2.2.2c) requires the integration of the variation equations of G and H [13, Knobloch]. We choose $\alpha = 0.1$, $\alpha = 0.01$, respectively in (1) and use equidistant partitions (5.1.1) in Algorithm 2 with $N_n^{max} = 64$, $N_n^{max} = 128$, respectively in Example 7.1, 7.2, respectively. The tolerance is set to $\varepsilon = 10^{-5}$. By choosing $k = 1$ and $k = -1$ in (4) the same invariant curve is computed. Thus, different choices $k \in \mathbf{Z}$ in (4) yield a test for the implementation. For summarizing the numerical experiments we adopt the notation $\max|\quad|$ introduced and used in Sections 6.2 and 6.3. Following the definition of the function b_n in (2.2.2d), $\max|b_n|$, $n = 0, 1, 2, \ldots.$ measures the accuracy with which the n^{th} Newton-Raphson approximation S_n satisfies the condition of invariance (1.1.7) in the knots

$$\varphi_j = \frac{\pi}{64} \cdot j, \; j = 0, \ldots, N_n^{max} .$$

The data generated by Algorithm 2 is given in Tables 16-19, and in the Figures 17 and 18 the invariant curves S are plotted.

Example 7.1: Let $\alpha = 0.01$, $c = 0.1$, $\nu = 1$ in (1) and $S_0(\varphi) = 1.8$ in Algorithm 1.

1. $T = -2\pi$ and $k = -1$ in (4)

N	$\min\|a_n\|$	$\max\|a_n\|$	$\max\|d_n\|$	$\max\|b_n\|$	$\min\|S_{n+1}\|$	$\max\|S_{n+1}\|$
0	1.044	1.046	$3.2 \cdot 10^{-1}$	$1.4 \cdot 10^{-2}$	1.966	2.118
1	1.062	1.077	$6.3 \cdot 10^{-2}$	$4.9 \cdot 10^{-3}$	1.946	2.054
2	1.060	1.070	$3.1 \cdot 10^{-3}$	$2.2 \cdot 10^{-4}$	1.945	2.052
3	1.060	1.070	$7.7 \cdot 10^{-6}$	$5.4 \cdot 10^{-7}$	1.945	2.051

Table 16

2. T = 2π and k = 1 in (4)

N	$\min\|a_n\|$	$\max\|a_n\|$	$\max\|d_n\|$	$\max\|b_n\|$	$\min\|S_{n+1}\|$	$\max\|S_{n+1}\|$
0	0.955	0.956	$3.1 \cdot 10^{-1}$	$1.3 \cdot 10^{-2}$	1.964	2.110
1	0.930	0.942	$5.6 \cdot 10^{-2}$	$3.9 \cdot 10^{-3}$	1.946	2.054
2	0.935	0.944	$2.1 \cdot 10^{-3}$	$1.4 \cdot 10^{-4}$	1.946	2.051

Table 17

Example 7.2: Let $\alpha = 0.1$, $c = 0.1$, $\nu = 1$ in (1) and $S_0(\varphi) = 1.9$ in Algorithm 1.

1. T = –2π and k = –1 in (4)

N	$\min\|a_n\|$	$\max\|a_n\|$	$\max\|d_n\|$	$\max\|b_n\|$	$\min\|S_{n+1}\|$	$\max\|S_{n+1}\|$
0	1.527	1.811	$2.8 \cdot 10^{-1}$	$1.5 \cdot 10^{-2}$	1.893	2.189
1	1.790	2.180	$7.0 \cdot 10^{-2}$	$8.3 \cdot 10^{-3}$	1.893	2.120
2	1.787	1.978	$7.6 \cdot 10^{-3}$	$7.5 \cdot 10^{-3}$	1.893	2.111
3	1.787	1.965	$8.0 \cdot 10^{-5}$	$7.7 \cdot 10^{-5}$	1.893	2.111

Table 18

2. T = 2π and k = 1 in (4)

N	$\min\|a_n\|$	$\max\|a_n\|$	$\max\|d_n\|$	$\max\|b_n\|$	$\min\|S_{n+1}\|$	$\max\|S_{n+1}\|$
0	0.541	0.602	$2.3 \cdot 10^{-1}$	$9.5 \cdot 10^{-2}$	1.893	2.132
1	0.503	0.559	$2.1 \cdot 10^{-2}$	$1.0 \cdot 10^{-2}$	1.893	2.111
2	0.509	0.559	$1.5 \cdot 10^{-4}$	$7.4 \cdot 10^{-4}$	1.893	2.111

Table 19

Van-der-Pol $c = 0.1$ $\varepsilon = 0.01$ $\nu = 1$

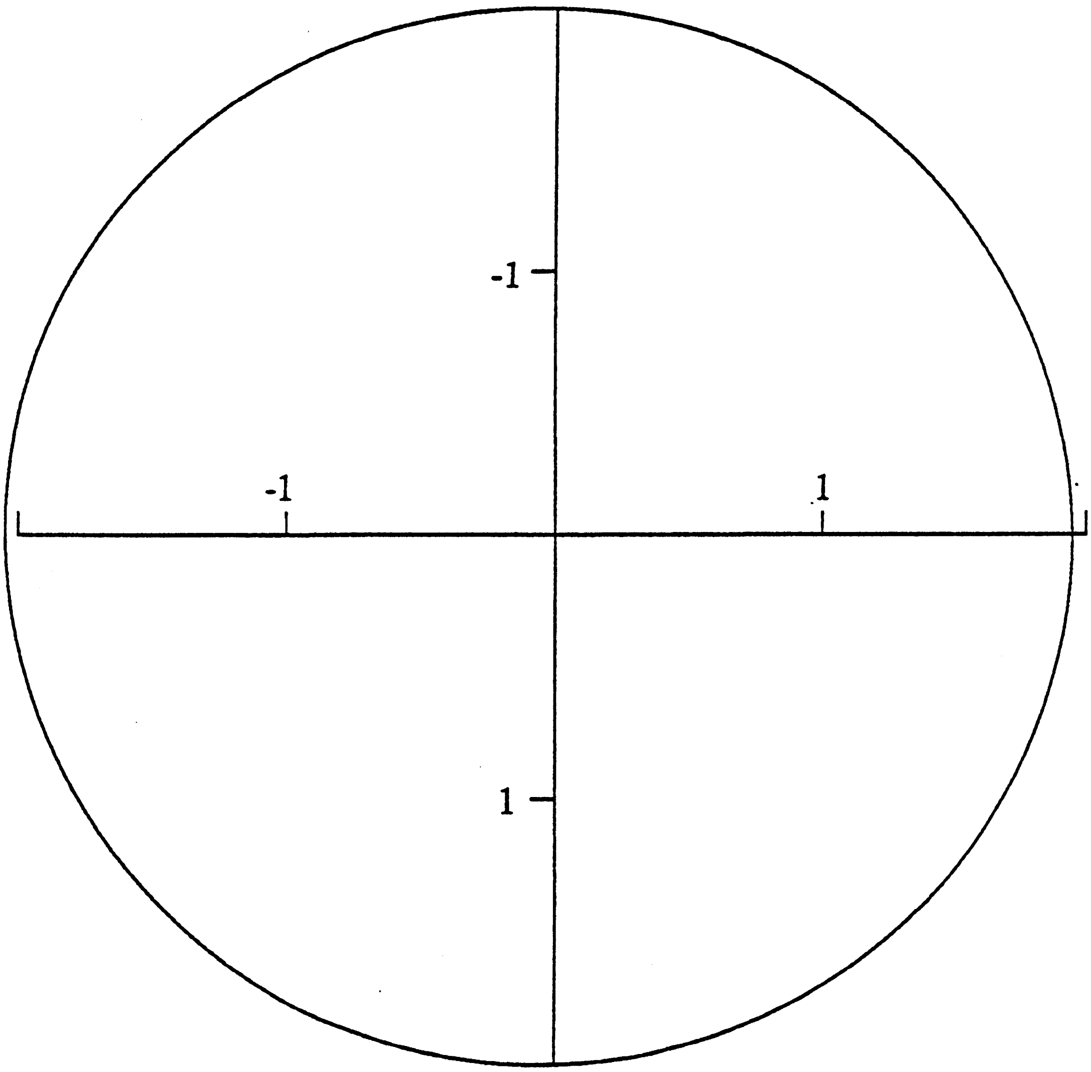

Figure 17

Van-der-Pol c = 0.1 ε = 0.1 ν = 1

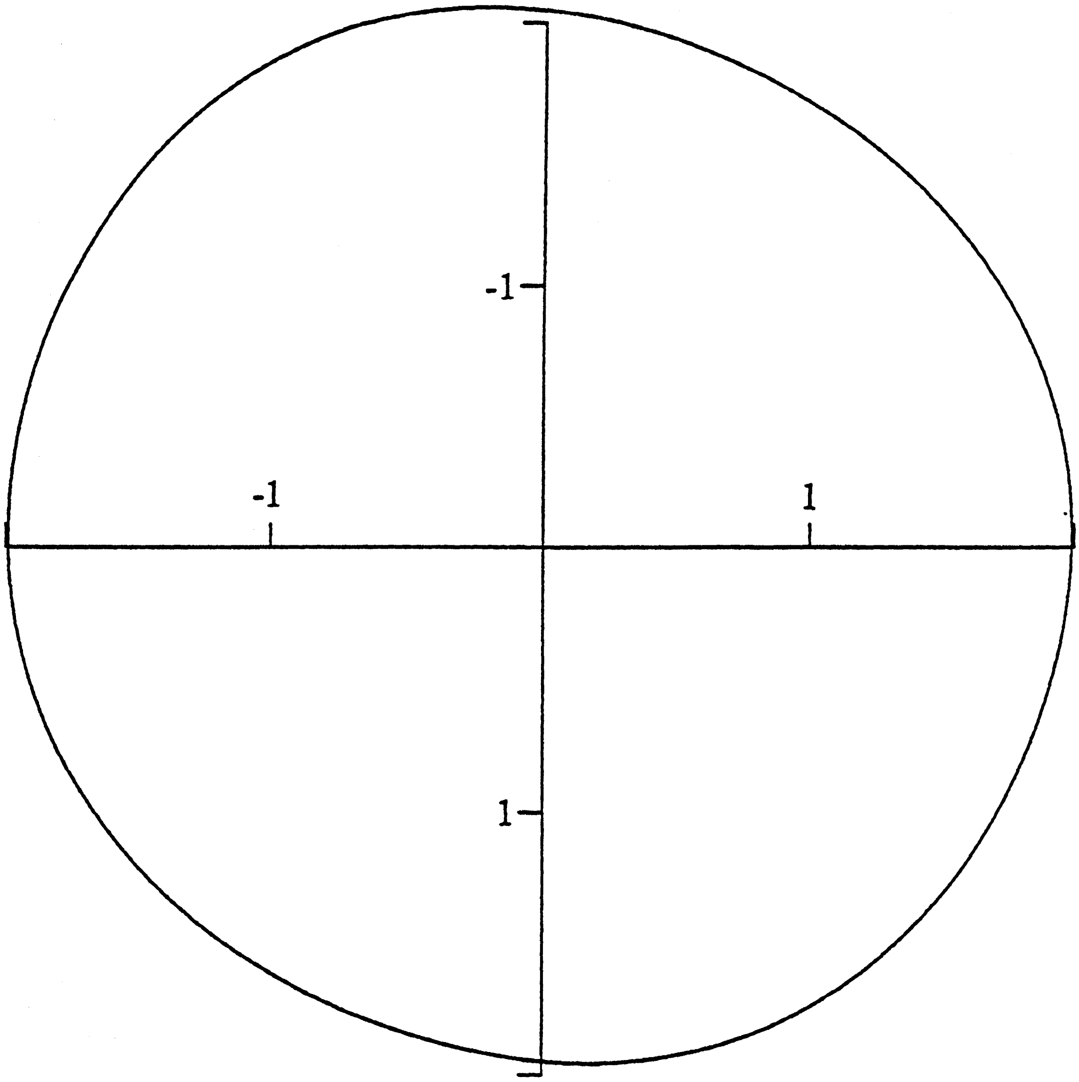

Figure 18

8. The numerical computation of the rotation number

8.1. Approximating the solution of the Abel equation

The following section is based on the considerations in Section 4.3 and Section 5.2. It has to be seen as a by-product of the Fourier Method (see Algorithm 3). We derive an algorithm for approximating the rotation number introduced in (4.3.2) and compare our calculation with the numerical evaluation of (4.3.2). There are other efficient approaches for calculating the rotation number of circle maps [29, Veldhuizen], [17, MacKay], and a comparison between the different approaches is left to future research.

Starting point is Lemma 4.7, i.e. it is assumed that the invertible circle map h: $\mathbf{R}^1 \to \mathbf{R}^1$

$$h(x) = x + p(x),\ p \in P,\ x \in \mathbf{R}^1 \tag{8.1.1}$$

has bounded variation and the rotation number ρ of h is irrational. Applying Lemma 4.7 and combining (4.3.38), (4.3.39) and (4.3.40) yield the *Abel equation*

$$(D \circ h)(x) - D(x) = 2\pi\rho \tag{8.1.2}$$

[16, Kuczma] where ρ is the rotation number of h and D: $\mathbf{R}^1 \to \mathbf{R}^1$ is an invertible function we represent by

$$D(x) = x + d(x),\ d \in P. \tag{8.1.3}$$

We are concerned with the numerical approximation of D in (3), (2), respectively and ρ in (2). The methodology developed in Section 5.2 is used.

Let $B_k(x)$, k = 0, 1, 2,…., $x \in \mathbf{R}^1$ be the basis of the real Fourier series, defined in (5.2.7). We consider the real Fourier series

$$p(x) = \sum_{k=0}^{\infty} h_k B_k(x) \tag{8.1.4}$$

with

$$h_0(x) = \frac{1}{2\pi}\int_0^{2\pi} p(x)dx$$

and

$$h_k(x) = \frac{1}{\pi}\int_0^{2\pi} p(x)B_k(x)dx\,,\ k = 1, 2,\ldots.$$

exists. In addition, it is assumed that there exists also real Fourier series for

$$L\,B_k(x) = (B_k \circ h)(x) - B_k(x),\ k = 0, 1, 2,\ldots, \tag{8.1.5}$$

i.e. there exists $a_{m,k} \in \mathbf{R}^1$ such that

$$LB_k(x) = \sum_{m=0}^{\infty} a_{m,k} B_k(x) \tag{8.1.6}$$

with

$$a_{0,k} = \frac{1}{2\pi}\int_0^{2\pi} LB_k(x)\,dx\,,\ k = 0, 1, 2,\ldots.$$

and

$$a_{m,k} = \frac{1}{\pi}\int_0^{2\pi} LB_k(x)\,B_m(x)\,dx\,,\ m = 1, 2,\ldots,\ k = 0, 1, 2,\ldots.$$

Using (1), (2) and (3) we have

$$(d \circ h)(x) - d(x) = 2\pi\rho - p(x). \tag{8.1.7}$$

With $f(x) = d(x)$, $a(x) = 1$, $b(x) = 2\pi\rho - p(x)$ it follows that (7) is of the form (2.3.2), thus the linear operator L in (2.3.3) is given by

$$L(d) = d \circ h - d. \tag{8.1.8}$$

Let

$$d(x) = \sum_{k=0}^{\infty} d_k B_k(x)\,.$$

Using (5), (6), (8) and applying the linearity of L it follows that

$$\sum_{k=0}^{\infty} d_k \sum_{m=0}^{\infty} a_{m,k} B_m(x) = 2\pi\rho - p(x).$$

For k = 0 in (5) and (5.2.9) we have $L(B_0) = L(1) = 0$, i.e. $a_{m,0} = 0$, m = 0, 1, 2, …. Equating coefficients for B_k, k = 0, 1, 2,…. yields the infinite system of linear equations

$$\begin{bmatrix} 0 & a_{0,1} & a_{0,2} & a_{0,3} & a_{0,4} & \cdot \\ 0 & a_{1,1} & a_{1,2} & a_{1,3} & \cdot & \cdot \\ 0 & a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot \\ 0 & a_{3,1} & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ d_3 \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} 2\pi\rho - h_0 \\ h_1 \\ h_2 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \qquad (8.1.9)$$

for the unknowns (d_0, d_1, d_2, d_3,....). The first equation in (9) is the same as

$$2\pi\rho = h_0 + \sum_{k=1}^{\infty} a_{0,k} d_k \qquad (8.1.10)$$

and it is seen that $d_0 \in \mathbf{R}^1$ is arbitrary. We normalise with $d_0 = 0$ (compare with Lemma 4.7) and consider the infinite system

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & \cdot & \cdot \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdot & \cdot & \cdot \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdot & \cdot & \cdot \\ a_{4,1} & a_{3,3} & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \qquad (8.1.11)$$

for the unknown coefficients (d_1, d_2, d_3,….). In the following (10) and (11) are used for the approximation of ρ. In order to approximate the coefficients $a_{m,k}$, m = 0, 1,…., N − 1, k = 1,…., N − 1 in (10) and (11) we compute the right side of (8) for k = 1,…., N − 1 in the knots

$$x_j = \frac{2\pi}{N} \cdot j,\ j = 0,…., N-1,\ N = 2^i,\ i = 3, 4,…. \qquad (8.1.12)$$

and use the Fast Fourier Transformation (F.F.T.) for transforming to real Fourier coefficients. We solve the finite system

$$\hat{A}\,\hat{d} = \hat{h},\ \hat{A} = \begin{bmatrix} \hat{a}_{0,0} & \hat{a}_{0,1} & \hat{a}_{0,2} & \cdot & \cdot & \hat{a}_{0,N-1} \\ \hat{a}_{1,0} & \hat{a}_{1,1} & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \hat{a}_{N-1,0} & \cdot & \cdot & \cdot & \cdot & \hat{a}_{N-1,N-1} \end{bmatrix},\ \hat{h} = \begin{bmatrix} \hat{h}_0 \\ \hat{h}_1 \\ \cdot \\ \cdot \\ \cdot \\ \hat{h}_{N-1} \end{bmatrix} \tag{8.1.13}$$

for the unknowns $(\hat{d}_1, \hat{d}_2, \ldots, \hat{d}_{N-1})$ and approximate ρ by

$$2\pi\rho_N = h_0 + \sum_{k=1}^{N-1} \hat{a}_{0,k}\hat{d}_k\,. \tag{8.1.14}$$

By increasing N of (13) and (14) we investigate the convergence of ρ_N. For the numerical experiments we choose

$$h_0 = 1,\ 0 \le h_1 \le 1,\ h_k = 0,\ k \ge 2$$

in (4), i.e. with $\alpha = h_1$ in (4) we have

$$p(x) = 1 + \alpha \sin x. \tag{8.1.15}$$

By using complex Fourier series

$$D(x) = x + \sum_{k=-\infty}^{\infty} \delta_k e^{ikx}$$

with $\delta_0 = 0$ (compare Lemma 4.7) we find that the coefficients δ_k, $|k| \in \mathbf{N}$ are the solution of the infinite system of linear equations

$$\begin{pmatrix} & & \cdot & & \\ & & \cdot & & \\ & & \cdot & & \\ & e^{-ih_0}J_0(-h_1)-1 & 0 & e^{-ih_0}J_{-2}(h_1) & \\ \cdot\ \cdot\ \cdot & e^{-ih_0}J_1(-h_1) & 0 & e^{-ih_0}J_{-1}(h_1) & \cdot\ \cdot\ \cdot \\ & e^{-ih_0}J_2(-h_1) & 0 & e^{-ih_0}J_0(h_1)-1 & \\ & & 0 & & \\ & & \cdot & & \\ & & \cdot & & \end{pmatrix} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \delta_1 \\ 0 \\ \delta_{-1} \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} \\ \\ \cdot \\ 0.5i\,h_1 \\ 0 \\ -0.5i\,h_1 \\ \cdot \\ \\ \\ \end{pmatrix}$$

and for complex Fourier series, (10) is given by

$$2\pi\rho = h_0 + \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \delta_k e^{ikh_0} J_{-k}(kh_1),$$

where J_k denotes the k^{th} Besselfunction (compare Example 5.1). As described in Section 5.2 real Fourier series are more efficient for the numerical experiments.

Gauss elimination is used for solving the systems of linear equation. In Table 20 we give the approximations ρ_N, $N = 2^i$, $3 \le i \le 9$ for ρ. The computer times are in the last line of Table 20. We observe that they depend critically on N and vary little for different choices of α.

α	ρ_8	ρ_{16}	ρ_{32}
0.2	0.15620'40786'814	0.15620'40802'397	0.15620'40802'913
0.4	0.14686'00621'090	0.14686'07310'353	0.14686'07108'958
0.6	0.12922'98691'699	0.12926'07661'144	0.12926'15963'618
0.8	0.09721'44994'303	0.09790'75391'843	0.09793'82258'314
[sec]	10^{-3}	$5 \cdot 10^{-3}$	$2 \cdot 10^{-3}$

α	ρ_{64}	ρ_{128}	ρ_{256}
0.2	0.15620'40802'913	0.15620'40802'913	0.15620'40802'913
0.4	0.14686'07108'685	0.14686'07108'685	0.14686'07108'684
0.6	0.12926'10492'661	0.12926'10495'272	0.12926'10495'271
0.8	0.09801'02779'331	0.09794'99842'359	0.09795'01138'223
[sec]	10^{-1}	$7 \cdot 10^{-1}$	5

α	ρ_{512}
0.2	0.15620'40802'912
0.4	0.14686'07108'684
0.6	0.12926'10495'270
0.8	0.09795'01138'260
[sec]	36

Table 20

For comparing the figures in Table 20 we choose $x_0 \in \mathbf{R}^1$ and compute the approximations

$$\frac{h^k(x_0)}{2\pi k},\ k = 10^i,\ 3 \le i \le 6. \tag{8.1.16}$$

Table 21 contains the numerical results for $x_0 = 0$ and $x_0 = 1$ in (16).

$x_0 = 0$

α	1000	10000	100000	1000000
0.2	0.15621'38638	0.15620'38783	0.15620'40614	0.15620'40781
0.4	0.14688'87662	0.14687'20503	0.14686'07046	0.14686'07892
0.6	0.12940'38773	0.12927'59409	0.12926'23433	0.12926'10536
0.8	0.09795'16740	0.09797'42912	0.09795'01300	0.09795'01935

$x_0 = 1$

α	1000	10000	100000	1000000
0.2	0.15621’38638	0.15622'05618	0.15620'62827	0.15620'42523
0.4	0.14688'87662	0.14688'17086	0.14686'25689	0.14686'08867
0.6	0.12961'72630	0.12928'23237	0.12926'34944	0.12926'12443
0.8	0.09721'44994	0.09790'75391	0.09795'18255	0.09795'47215
[sec]	$1.7 \cdot 10^{-1}$	$1.6 \cdot 10^{-1}$	2.2	2.4

Table 21

The computer time for the calculation of the approximation of ρ is given in the last line of Table 21. Further initial values $x_0 \in \mathbf{R}^1$ in (16) show that the number of correct digits are little dependent on ρ and x_0.

The comparison of the data shows that the method for computing ρ derived in this section is more efficient and more precise than evaluating (16). By doubling the knots (12) the convergence is observed to be quadratic. The values for ρ_{512} in Table 20 are computed with machine precision. We use them as reference values for the values in Table 21 and find empirically

$$\left| \rho - \frac{h^k(x_0)}{2\pi k} \right| \approx \frac{1}{k},\ k = 10^i,\ 3 \le i \le 6.$$

8.2. Intervals with rational rotation number

In this section we investigate the rotation number ρ of (8.1.1) with

$$p(x) = 1 + \alpha \sin x,\ \alpha \in [0, 1],\ x \in \mathbf{R}^1 \qquad (8.2.1)$$

as a function of α, i.e. the function ρ(α) is considered. ρ(α) is monotonically decreasing [9, Herman] with $\rho(0) = \frac{1}{2\pi}$, ρ(1) = 0 and there exists nonempty intervals [a, b] ⊂ [0, 1], in which $\rho(\alpha) \in [0, \frac{1}{2\pi}]$ is rational and constant.

We start by describing the computation of a and b. Let $\rho = \frac{p}{q} \in \mathbf{Q}$ be in $[0, \frac{1}{2\pi}]$ with p ∈ **N**, q ∈ **N** non divisible. It follows that the map

$$P(x, \alpha) = h^q(x) - 2\pi p,\ \alpha \in (a, b) \qquad (8.2.2)$$

has 2q fixed points in the first argument. If α = a, α = b, respectively 2 fixed point are identical and P(x, α) has q double fixed points. Thus, a and b are two different solutions for α of the nonlinear system of the equations

$$P(x, \alpha) = x$$

$$\frac{d}{dx} P(x, \alpha) = 1. \qquad (8.2.3)$$

(3) cannot be solved analytically and we are bound to use a numerical solver. The right side of (3) needs to be evaluated for any given argument x and α. With $h^0(x) = x$, $\frac{d}{dx} h^0(x) = 1$ we use for n = 1, 2 , 3,…. the recursions

$$h^n(x) = h(h^{n-1}(x)) = h^{n-1}(x) + 1 + \alpha \sin(h^{n-1}(x)),$$

$$\frac{dh^n(x)}{dx} = h'(h^{n-1}(x)) \frac{dh^{n-1}(x)}{dx^{n-1}} = (1 + \alpha \cos(h^{n-1}(x))) \frac{dh^{n-1}(x)}{dx^{n-1}}.$$

Different initial values for solving (3) yield different solutions of (3) and a, b can be computed. We have computed many intervals. They have different lengths and the numerical experiments show that the maximum length d = b – a is approximately $d \approx 10^{-3}$. In view of the following consideration we note in Table 22 a selection of intervals with decreasing length d.

a	b	d	p	q
0.76322'68758'1354	0.76333'37295'8432	10^{-4}	2	19
0.77277'58332'3842	0.77278'41847'8643	10^{-6}	3	29
0.71488'95568'8848	0.71488'95718'5987	10^{-8}	5	44
0.78494 '07989'3481	0.78494'07990'3550	10^{-10}	10	99

Table 22

The method derived in Section 8.1 is applied for computing the rotation number ρ, i.e. we choose $\alpha \in [a, b]$, compute the solution of the linear system (8.1.13) for $N = 2^i$, $3 \le i \le 9$ and observe the convergence ρ_N in (8.1.14). Table 23 shows the results of the computation.

$\alpha = \frac{b - a}{2}$	ρ^{-1}	ρ_8^{-1}	ρ_{16}^{-1}
0.76328'03026'989	9.5	9.53752'05054'620	9.50124'91733'634
0.77278'00901'243	9.666…	9.71078'29568'092	9.66745'58733'453
0.71488'95643'742	8.8	8.81489'28478'860	8.79997'16662'026
0.78494'07989'852	9.9	9.95211'22783'922	9.89758'50075'933

ρ_{32}^{-1}	ρ_{64}^{-1}	ρ_{128}^{-1}
9.49899'70203'605	9.50001'25123'950	9.50020'41817'237
9.66674'28979'902	9.66674'28979'880	9.66666'65011'520
8.79818'62967'682	8.79998'99976'826	8.80000'00035'632
9.89485'30092'630	9.89642'84916'976	9.90003 '73677'284

ρ_{256}^{-1}	ρ_{512}^{-1}
9.50097'45349'552	9.49922'52350'300
9.66678'90405'020	9.66665'18787'454
8.80000'00000'038	8.79992'20026'098
9.90000'00021'863	9.90000'00000'106

Table 23

From Table 23 it is seen that that ρ_N, $N = 2^i$, $3 \leq i \leq 9$ does not converges towards ρ. However, N being sufficiently large, the number of exact digits increases with decreasing length of the interval d.

If ρ_N converges quadratically it cannot be decided whether the rotation number ρ is irrational or ρ is in an interval [a, b] with a length equal or smaller than the machine precision. In order to proof this conjecture we work with 25 digit precision and compute 3 very short intervals. The numerical evidence is given in Table 25. In addition, we note that the values in Tables 24 and 25 yield a test of our implementation.

ρ	a	b
$\frac{6}{43}$	0.49725'75759'80532'81412'99710	0.49725'75759'80546'21358'1287
$\frac{10}{89}$	0.72268'91065'60072'68085'61050	0.72268'91065'61834'80508'4647
$\frac{20}{199}$	0.78723'96808'64321'62408'80481	0.78723'96808'64365'58229'91089

Table 24

α	ρ^{-1}	ρ_4^{-1}	ρ_8^{-1}
0.49725'75759'8054	7.1666	7.33156'43541'65	7.16694'01180'76
0.72268'91065'6100	8.9	10.19905'18044'81	8.91630'40295'55
0.78723'96808'6433	9.9555	12.36268'18556'49	10.00045'60023'72

ρ_{16}^{-1}	ρ_{32}^{-1}	ρ_{64}^{-1}
7.16668'25592'01	7.16666'59896'75	7.16666'66666'69
8.89910'98441'07	8.89731'15634'25	8.89992'70420'11
9.94348'71039'56	9.94069'18619'50	9.94349'03392'50

ρ_{128}^{-1}	ρ_{256}^{-1}	ρ_{512}^{-1}
7.16666'66666'67	7.16666'66666'67	7.16666'66666'67
8.89999'98670'73	8.89999'99999'98	8.90000'00000'00
9.94759'37092'32	9.94999'97347'88	9.94999'99999'99

Table 25

9. The model of delayed regulation

9.1. Numerical results from the Fourier method

Following Section 1.2 we consider in this chapter the model of *delayed regulation*

$$x_{n+1} = \gamma\, x_n \cdot (1 - x_{n-1}),\ n = 0, 1, 2, ..., \gamma \in \mathbf{R^1} \tag{9.1.1}$$

where $x_0 \in \mathbf{R^1}$ is assumed to be a given initial value. The iteration (1) is studied in biology [19, Maynard Smith]. In Veldhuizen [27] and [28] spline methods are applied to (1). The computations described in this chapter are based on real Fourier series (see Algorithm 3 in Section 5.2).

With $x \in \mathbf{R^1}$, $y \in \mathbf{R^1}$ as origins and $\tilde{x} \in \mathbf{R^1}$, $\tilde{y} \in \mathbf{R^1}$ as images the map (1) is equivalent to the map $\mathbf{R^2} \to \mathbf{R^2}$ defined by

$$\begin{aligned} \tilde{x} &= y \\ \tilde{y} &= \gamma\, y \cdot (1 - x). \end{aligned} \tag{9.1.2}$$

For $1 < \gamma \le 2$, the map (2) has a stable fixed point

$$x = y = 1 - \frac{1}{\gamma}. \tag{9.1.3}$$

If $\gamma \in (2, 2.27)$ the fixed point (3) is unstable and there exists a simple closed attractive invariant curve S (see Definition 1.5) of (2) that surrounds the fixed point (3) [2, Aronson].

In order to satisfy Assumption 1.3 we choose the origin of the map (2) in (3) and transform to polar coordinates. We find

$$\begin{aligned} \tilde{x} &= y \\ \tilde{y} &= x + y - \gamma\, x \cdot (1 + y) \end{aligned} \tag{9.1.4}$$

and with

$$x = r \cos \varphi,\ y = r \sin \varphi,\ \tilde{x} = \tilde{r} \cos \tilde{\varphi},\ \tilde{y} = \tilde{r} \sin \tilde{\varphi}$$

we follow (1.1.4) by introducing the maps G and H:

$$\tilde{r} = G(r, \varphi)$$

$$\tilde{\varphi} = H(r, \varphi)$$

where G and H are the given by

$$G(r, \varphi) = \sqrt{(r\sin\varphi)^2 + (g(r,\varphi))^2}$$

$$H(r, \varphi) = \arctan\frac{g(r,\varphi)}{r\sin\varphi}$$

with

$$g(r, \varphi) = r\cos\varphi + r\sin\varphi - \gamma\, r\cos\varphi \cdot (1 + r\sin\varphi).$$

The uniqueness of H is ensured by the choice $\varphi \in [0, 2\pi)$. In the following we consider the iteration

$$\begin{aligned} r_{n+1} &= G(r_n, \varphi_n) \\ \varphi_{n+1} &= H(r_n, \varphi_n) \end{aligned} \qquad (9.1.5)$$

where n = 0, 1, 2,…. and the initial values $r_0 \in \mathbf{R}^1$, $\varphi_0 \in \mathbf{R}^1$ are assumed to be given.

As S is attractive it follows that the iteration (5) of a point $r_0 \in \mathbf{R}^1$, $\varphi_0 \in \mathbf{R}^1$ that is in a sufficiently small annular neighbourhood of S converges towards elements of S. With 680 iterations we find 16 different points with polar coordinates (R_j, φ_j), $0 \le j \le 15$ with

$$\left|\varphi_j - \frac{\pi}{8}\cdot j\right| < 10^{-2},\ j = 0, 1, 2, \ldots, 15.$$

In Table 26 we note the numerical values for R_j and φ_j, j = 0, 1, 2,…., 15.

j	0	1	2	3	4	5	6	7
R_j	0.27	0.30	0.33	0.30	0.23	0.18	0.17	0.17
j	8	9	10	11	12	13	14	15
R_j	0.21	0.3	0.43	.35	0.30	0.28	0.27	0.27

Table 26

The data in Table 26 is used for the initial approximation in Algorithm 1 (see Section 2.2). We compute the Fourier series of S for the parameters γ = 2.10 and γ = 2.11 in (2) by solving the system (5.2.15) for N = 64, N = 128, respectively. Let $\tilde{\delta}_n$, δ_n, respectively denote the solution for N = 64, N = 128, respectively in the n^{th} Newton-Raphson step, n = 0, 1, 2. In Table 27 we give the norms $\|\tilde{\delta}_n - \delta_n\|_{max}$ and $\|\delta_n\|_{max}$ in Algorithm 3.

	$\gamma = 2.10$		$\gamma = 2.11$	
n	$\|\tilde{\delta}_n - \delta_n\|_{max}$	$\|\delta_n\|_{max}$	$\|\tilde{\delta}_n - \delta_n\|_{max}$	$\|\delta_n\|_{max}$
0	$1 \cdot 10^{-5}$	$2 \cdot 10^{-2}$	$2 \cdot 10^{-6}$	$1 \cdot 10^{-2}$
1	$1 \cdot 10^{-5}$	$2 \cdot 10^{-4}$	$6 \cdot 10^{-6}$	$7 \cdot 10^{-4}$
2	$1 \cdot 10^{-6}$	$4 \cdot 10^{-6}$	$1 \cdot 10^{-6}$	$3 \cdot 10^{-5}$

Table 27

In order to check the tolerance of the computation we iterate the point (S(0), 0) and test the condition of invariance (1.1.7) in 100 points of S. For both values γ = 2.10 and γ = 2.11 the condition of invariance (1.1.7) is satisfied for 7 digit precision (compare the error analysis for γ = 2.11 in Table 2 of [28]). The graphs for γ = 2.10, γ = 2.11, respectively are in Figure 19, 20, respectively.

9.2. An example with rational rotation number

Following (9.1.4), the map F: $\mathbf{R}^2 \rightarrow \mathbf{R}^2$ in (1.2.2) is given by

$$F:\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} y \\ x+y-\gamma\, x\ (1+y) \end{pmatrix}.$$

The rotation number of the invariant curve S under F as a function of γ is depicted in Aronson [2]. There exists intervals with phase locking, i.e. intervals in which the rotation number is rational and constant. We start with the description of their computation. Let

$$F^{n+1}(x, y) = F(F^n(x, y)),\ n = 1, 2, \ldots. \tag{9.2.1}$$

denote the n^{th} iterate of F where

$$F^1 = F.$$

If $\rho = \frac{p}{q}$, there exists $x_0 \in \mathbf{R}^1$, $y_0 \in \mathbf{R}^1$ with

$$F^q(x_0, y_0) = (x_0, y_0). \tag{9.2.2}$$

We denote with J^n, n = 0, 1, 2,…. the Jacobian of F^n. If $\lambda = 1$ is an eigenvalue of $J^q(x_0, y_0)$ the corresponding value γ of determines the beginning or the end of an interval with phase locking. Thus, together we (2) we consider

$$\mathrm{Det}(J^q(x_0, y_0) - I) = 0 \tag{9.2.3}$$

where

$$I = \begin{pmatrix} x_0 & 0 \\ 0 & y_0 \end{pmatrix}.$$

(2) and (3) define a system of nonlinear equations for the unknown $x_0 \in \mathbf{R}^1$, $y_0 \in \mathbf{R}^1$, $\gamma \in \mathbf{R}^1$. The system (2), (3) cannot be solved analytically and a numerical solver has to be used instead. Different solutions for (2) and (3)

yield intervals with phase locking. For evaluating the right side of (3) we use the recursions

$$J^n(x, y) = J(F^{n-1}(x, y)) \cdot J^{n-1}(x, y),\ n = 2, 3, \ldots.$$

where

$$J(x, y) = J^1(x, y) = \begin{pmatrix} 0 & 1 \\ 1-\gamma(1+y) & 1-\gamma x \end{pmatrix}.$$

In the following we consider the case $\rho = \frac{2}{13}$ and with 10 digits precision we find

$$I = [2.11757'33497,\ 2.11777'03479].$$

For γ in the interior of I and q = 13, equation (2) has 26 different solutions. We denote them with (s_n, s_{n+1}), $0 \le n \le 12$ and (t_n, t_{n+1}), $0 \le n \le 12$ where $s_{13} = s_0$ and $t_{13} = t_0$. For their computation we use a 2-dimensional Newton-Raphson method for maps $\mathbf{R}^2 \to \mathbf{R}^2$. The numerical values for $\gamma = 2.1176718$ are in Table 28.

n	s_n	t_n
0	0.2264599115	0.2709055203
1	0.2718568797	0.2138197352
2	-0.1116242558	-0.2116291594
3	-0.3512080979	-0.3547834560
4	-0.3094687584	-0.2772519031
5	-0.1470991808	-0.0890243922
6	0.1023829203	0.1685830256
7	0.2986843115	0.2998649124
8	0.1194952800	0.0043919841
9	-0.2899177483	-0.3335470080
10	-0.3501098966	-0.3353535519
11	-0.2410275392	-0.1994325311
12	-0.0284221803	-0.0337514403

Table 28

The sequence s_n, $0 \le n \le 12$ is a *stable orbit of F* because the iteration (1) of each sequence that starts with $x_0 \in \mathbf{R^1}$, $y_0 \in \mathbf{R^1}$ sufficiently close to S converges towards s_n, $0 \le n \le 12$. The sequence t_n, $0 \le n \le 12$ is an *unstable orbit of F* because of its stability under iteration of the inverse F^{-1}. Starting with two different initial values close of the instable orbit, Figures 22 and 23 show the evaluation of the iteration (1).

The computed Fourier series allows to approximate arbitrary points on S. In Figure 21 the invariant curve S is plotted. The iteration (1) cannot be used for this purpose. Applying the iteration (1) it is not clear which points on S are approximated with which precision. The values in Table 28 yield a check of our implementation:

1. We iterate (S(0), 0) with (1). The condition of invariance (1.1.7) is satisfied with 10 digit precision.

2. Let (S_n, φ_n), $0 \le n \le 12$ denote the polar coordinates of the values in Table 28. We interpolate the computed real Fourier polynomial of S in φ_n and find

$$\left\| S(\varphi_n) - S_n \right\| \approx 10^{-10}, \; 0 \le n \le 12.$$

For the discussion of the regularity of S, we calculate the characteristic exponents of the 13[th] iterate of (1). The numerical values are

$$\mu_{s_n} = \frac{\log \lambda_1}{\log \lambda_2} = 145.07, \; \mu_{t_n} = \frac{\log \lambda_1}{\log \lambda_2} = -149.50, \; 0 \le n \le 12.$$

λ_1, λ_2 are the eigenvalues of J^{13}. In s_n, $0 \le n \le 12$, S is 145 times differentiable and S is analytic in t_n, $0 \le n \le 12$. Thus S has high regularity. In accordance with our numerical experiments the real Fourier polynomial of S converges fast. The parameter $\gamma = 2.1176718$ can be treated without problem by using Algorithm 3 (see Section 5.2).

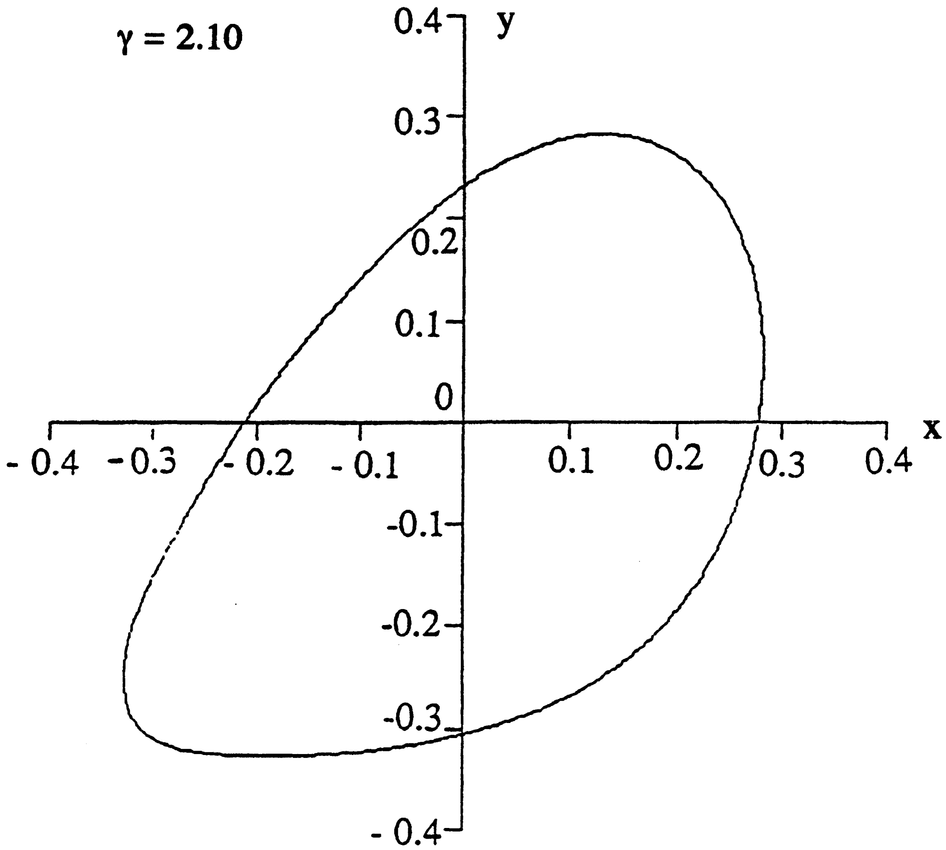

Figure 19

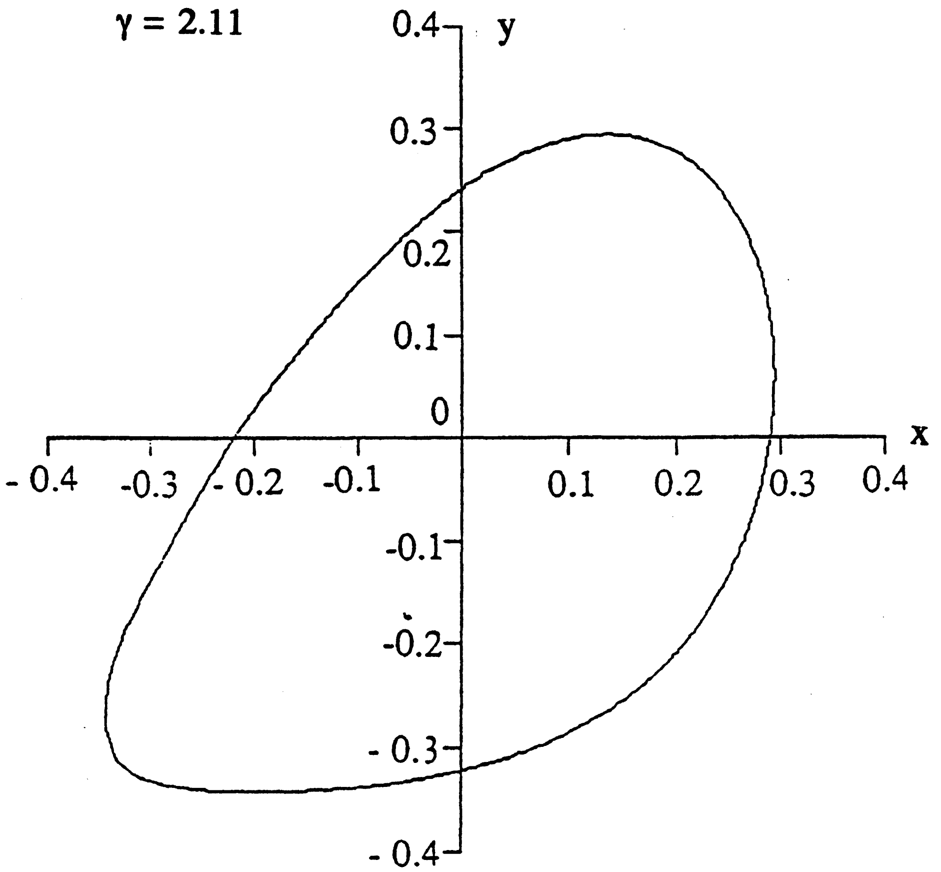

Figure 20

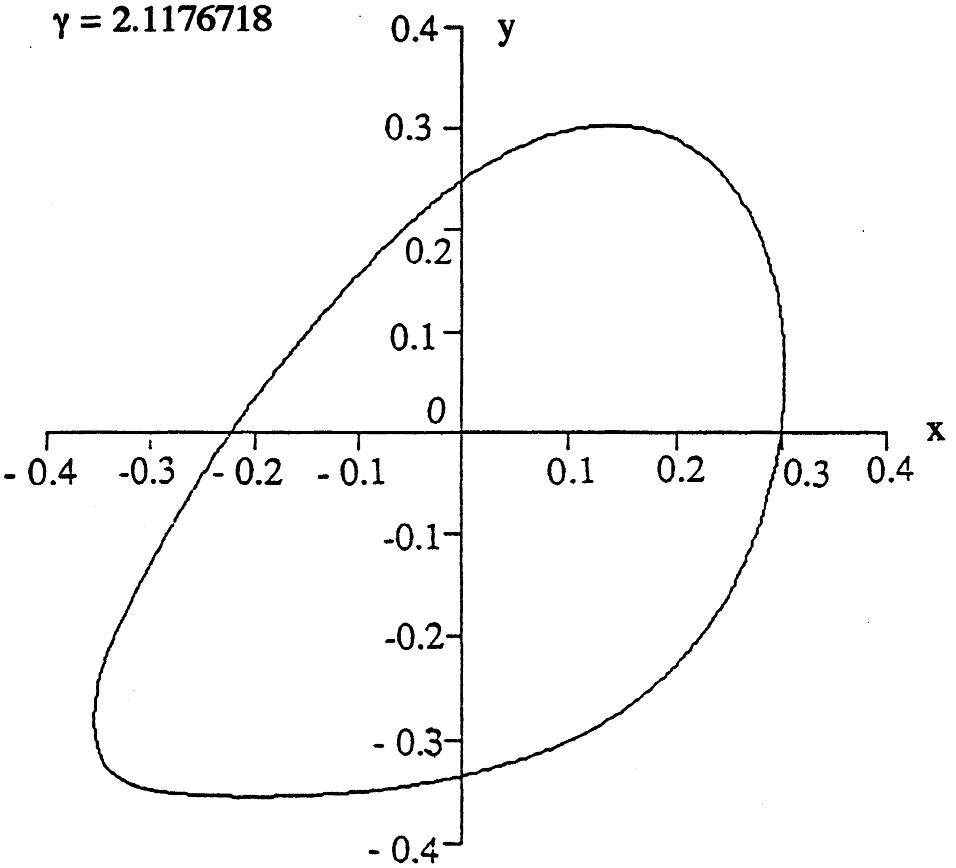

Figure 21

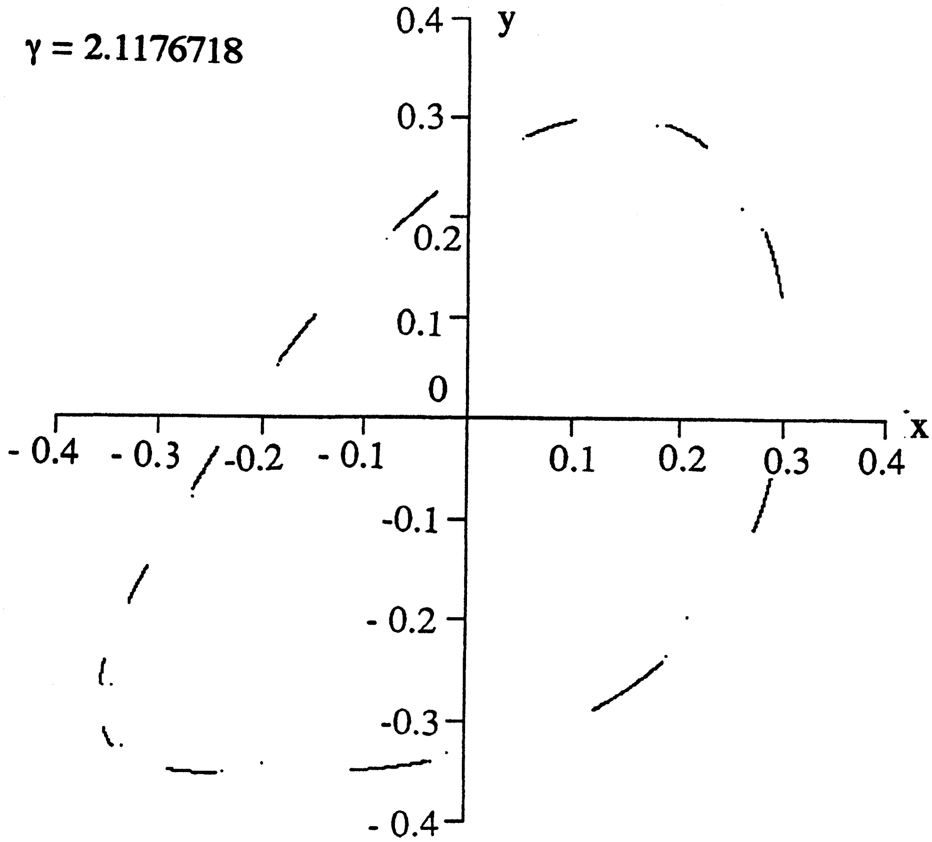

Figure 22

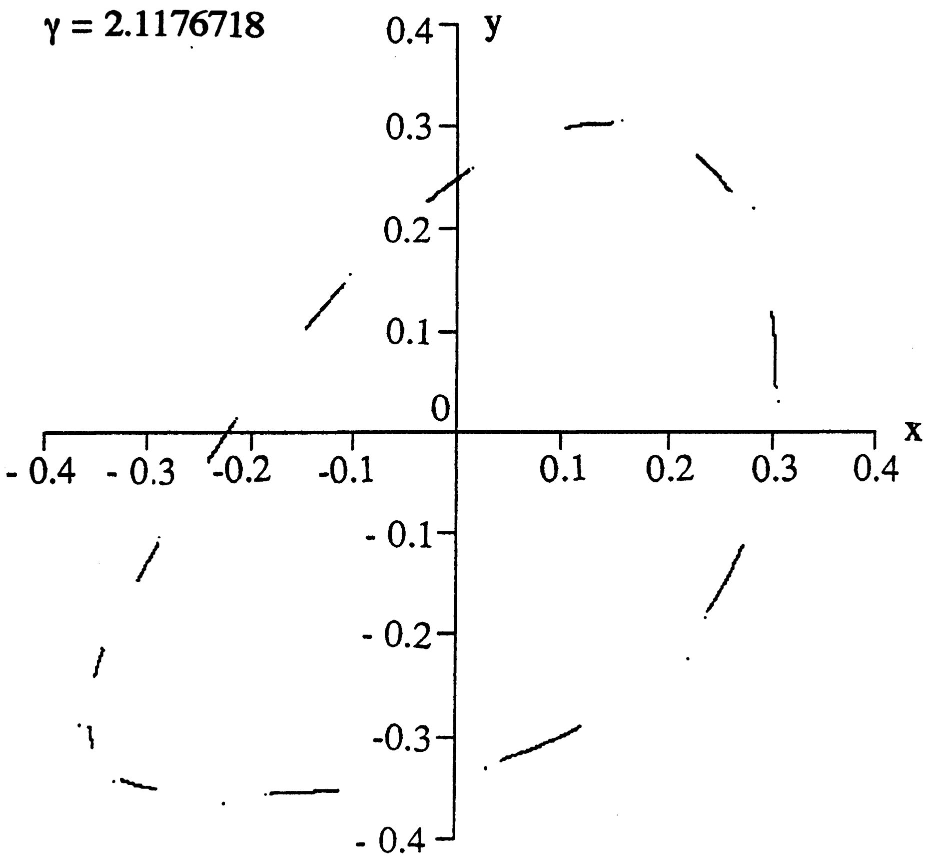

Figure 23

References

[1] Abramowitz M., Stegun A., 1965
Handbook of Mathematical Functions with formulas, graphs, and mathematical tables
Applied mathematics series, 55, 3rd printing

[2] Aronson D.G., Chary M.A., Hall G.R. and Mc Geehee R.P., 1982
Bifurcation from an Invariant Circle for Two-Parameter Families of Maps of the Plane: A Computer-Assisted Study
Communications in Mathematical Physics, 83, page 303-354

[3] Broer H.W., Osinga H.M., Vegter G., 1997
Algorithm for computing normally hyperbolic invariant manifolds
Zeitschrift für angewandte Mathematik und Physik, 48, page 480-534

[4] Chang T. N., 1983
Numerical Bifurcation Analysis of Simple Dynamical Systems
Dissertation, Concordia University of Montreal

[5] Eisenstat St., Elman H., Schultz M., 1983
Variational iterative methods for non symmetric systems of linear equations
SIAM J.Numer. Anal 20

[6] Guckenheimer J., and Holmes P., 1983
Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields
Springer, New York

[7] De Melo W. und Van Strien S., 1993
One - Dimensional Dynamics, 3.Folge, Band 25 der Reihe
Ergebnisse der Mathematik und ihrer Grenzgebiete
Springer, Berlin, Heidelberg

[8] Devaney, R., 2003
An Introduction to Chaotic Dynamical Systems
West view press

[9] Herman M. R., 1979
Sur la conjugaison différentiable des difféomorphismes du cercle a des rotations
lnstitut des hautes études scientifiques
Publications Mathématiques 49, page 5-233

[10] Kantorovitch L.W., Akilow G.P., 1981
Analyse Fonctionelle
Tome 1 et 2 Moscou, Editions Mir

[11] Kevrekidis I.G., Aris R., Schmidt L.D. and Pelikan S., 1985
Numerical computation of invariant circles of maps
Physica Vol. 16D, page 243-251

[12] Kirchgraber U., Stiefel E., 1978
Methoden der analytischen Störungsrechnung und ihre Anwendungen
Teubner Stuttgart

[13] Knobloch H.W., Kappel F., 1974
Gewöhnliche Differential-Gleichungen
Teubner Stuttgart

[14] Kreyzig E., 1978
Introductory Functional Analysis with Applications
John Wiley & sons

[15] Kuczma M., 1968
Functional equations in a single variable
Polish Scientific Publishers 46

[16] Kuczma M., Choczewski B., Ger R., 1990
Iterative functional equations
Encyclopedia of Mathematics and its Application 32 Cambridge University Press

[17] MacKay R.S., 1987
Rotation interval from a time series
J. Phys. A, 20:587-592

[18] Mäder R., 1980
Invariante Mannigfaltigkeiten
Diplomarbeit ETH Zürich

[19] Maynard Smith J., 1971
Mathematical ideas in biology
Cambridge University Press

[20] Nicolaisen N., 1998
Numerische Berechnung invarianter Kurven von Abbildungen
Dissertation, Universität Hamburg

[21] Osinga H., 1996
Computing Invariant Manifolds
Dissertation, Rijksuniversiteit Groningen

[22] Schumaker L., 1981
Spline functions basic theory
New York a.o. Wilney

[23] Stiefel E., 1955
Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssysteme
Comment. Math. Helv., 29

[24] Stoer J., 1983
Numerische Mathematik Band I und II
Springer

[25] Trubowitz E., Poeschel J., 1987
Inverse Spectraltheory
Pure and applied mathematics Academic Press

[26] Van-der-Pol B., 1927
Forced oscillations in a system with non-linear resistance
Phil. Mag.

[27] Van Veldhuizen M., November 1987
A new algorithm for the numerical approximation of an invariant curve
SIAM J.Sci. Stat. Comput. Vol. 8, No.6

[28] Van Veldhuizen M., October 1988
Convergence results for invariant curve algorithms
Mathematics of Computation
Vol. 51, Number 184, page 677-697

[29] Van Veldhuizen M., 1988
On the numerical approximation of the rotation number
Journal of Computational and Applied Mathematics 21 page 203-212

[30] Waldvogel J., Stenger F., Resch R., October 1996
Functional Equations Related to the Iteration of Functions
Research Report No. 96-14 ETH, CH-8092 Zurich

[31] Weyl H., 1916
Über die Gleichverteilung von Zahlen mod Eins
Mathematische Annalen 77, 1916 page 313